

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2004, 10:55:53 ; Search time 3023 Seconds  
(without alignments)  
9850.038 Million cell updates/sec

Title: US-09-976-673-11

Perfect score: 687  
Sequence: 1 atggtgagcgcgctgtctgaa.....tgcccgagaagccaactga 687

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenBank:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_cm:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vi:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rod:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	687	100.0	687	6	AX527898 Sequence
2	671	97.7	687	6	AX527900 Sequence
3	671	97.7	1396	6	AX527902 Sequence
4	669.4	97.4	1424	6	AX527904 Sequence
5	441.2	64.2	681	6	AX527894 Sequence
6	439.6	64.0	681	6	AX527896 Sequence
7	438.2	63.8	684	6	AX527910 Sequence
8	436.6	63.6	684	6	AX527892 Sequence
9	436.6	63.6	908	3	AF363776 Heteracti
10	435	63.3	910	6	AX527888 Sequence
11	435	63.3	910	6	AX527914 Sequence
12	431.8	62.9	908	6	AX527890 Sequence
13	429.2	62.5	680	6	AX527912 Sequence
14	424.8	61.8	699	12	AY233273 Synthetic
15	417.8	60.8	707	6	AX824739 Sequence
16	396.8	57.8	654	6	AX824741 Sequence
17	362.6	52.8	681	6	AX573118 Sequence
18	361.2	52.6	835	3	AF363775 Condy1act
19	361.2	52.6	835	6	AX573116 Sequence
20	358	52.1	684	3	AF383155 Condy1act
21	351.2	51.1	835	3	AY037777 Condy1act
22	306	44.5	678	12	AF506027 Synthetic
23	299	43.5	713	6	AX233627 Sequence
24	298.8	43.5	681	6	AX233584 Sequence
25	298.2	43.4	681	12	AF506025 Synthetic
26	298.2	43.4	1395	12	AF506026 Synthetic
27	295	42.9	10141	12	AY342347 Red H-Pel
28	294	42.8	10276	12	AY342348 Red H-Sti
29	292.4	42.6	678	6	AX370404 Sequence
30	292.4	42.6	678	6	AX370406 Sequence
31	292.4	42.6	678	6	AX824725 Sequence
32	291.4	42.4	4692	6	AX463702 Sequence
33	291.4	42.4	9320	6	AX663075 Sequence
34	290.8	42.3	6893	6	AX823860 Sequence
35	289.2	42.1	678	6	AX370408 Sequence
36	289.2	42.1	678	6	AX824732 Sequence
37	288.4	42.0	678	6	AX824729 Sequence
38	288.2	42.0	1050	6	AX666133 Sequence
39	287.6	41.9	675	6	AX824731 Sequence
40	268	39.0	955	3	AF246709 Anemonia
41	266.4	38.8	699	3	AF322222 Anemonia
42	264.8	38.5	696	6	AX686890 Sequence
43	264.8	38.5	696	6	AX824727 Sequence
44	261.2	38.0	720	6	AX666128 Sequence
45	261.2	38.0	9258	12	AF383623 Cloning v

ALIGNMENTS

RESULT 1  
AX527898 687 bp DNA linear PAT 21-NOV-2002  
LOCUS AX527898  
DEFINITION Sequence 11 from Patent WO0230965.  
ACCESSION AX527898  
VERSION AX527898.1 GI:25172348  
KEYWORDS  
SOURCE Heteractis crispa (leathery sea anemone)  
ORGANISM Heteractis crispa  
Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;  
Stichodactylidae; Heteractis.  
REFERENCE 1  
AUTHORS Lukyanov,S.A., Pradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.  
TITLE Nucleic acids encoding stichodactylidae chromoproteins  
JOURNAL Patent: WO 0230965-A 11 18-APR-2002;

FEATURES	source	location/Qualifiers
ORIGIN	Clontech Laboratories Inc. (US)	1..687 /organism="Heteractis crispa" /mol_type="unassigned DNA" /db_xref="taxon:175771"
Query Match	100.0%;	Score 687; DB 6; Length 687;
Best Local Similarity	100.0%;	Pred. No. 2.2e-79;
Matches 687;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 ATGTGAGCGGCTGCTGAAGGAGATATGCGCATCAAGATGTACATGAGGGCACCCTG	60
DB	1 ATGTGAGCGGCTGCTGAAGGAGATATGCGCATCAAGATGTACATGAGGGCACCCTG	60
QY	61 AAGCGCACTACTTCAAGTGCAGGGGCGAGGGCGCAGGGCAACCCCTTGCCTGCGCACCAG	120
DB	61 AAGCGCACTACTTCAAGTGCAGGGGCGAGGGCGCAGGGCAACCCCTTGCCTGCGCACCAG	120
QY	121 AGCATGAGAAATCCACGTGACCGAGGGGCGCCCTTGCCTTGCATCTCTGGCC	180
DB	121 AGCATGAGAAATCCACGTGACCGAGGGGCGCCCTTGCCTTGCATCTCTGGCC	180
QY	181 CCCTGCTGGAGTACCGCAGGAGACCTTCTGTGACCAACCGCGGAGATCCCGACTTC	240
DB	181 CCCTGCTGGAGTACCGCAGGAGACCTTCTGTGACCAACCGCGGAGATCCCGACTTC	240
QY	241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGAGAGAAACCAACCACTACGAGACGGC	300
DB	241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGAGAGAAACCAACCACTACGAGACGGC	300
QY	301 GGCATCCTGACCGCCACAGACAACGCTTGAGGGCAACTGCTGATCTTCAAGGTG	360
DB	301 GGCATCCTGACCGCCACAGACAACGCTTGAGGGCAACTGCTGATCTTCAAGGTG	360
QY	361 AAGGTGCAGCGGCAACCACTTCCCGCGGACGGCCCGGTGATGAAGAACAAGCGCGGC	420
DB	361 AAGGTGCAGCGGCAACCACTTCCCGCGGACGGCCCGGTGATGAAGAACAAGCGCGGC	420
QY	421 TGGAGGCCAGCAACCGAGGTGTGTACCCCGAGAACGGCGCTGTGCGCGCCGGAACGTG	480
DB	421 TGGAGGCCAGCAACCGAGGTGTGTACCCCGAGAACGGCGCTGTGCGCGCCGGAACGTG	480
QY	481 ATGGCCCTGAAGGTGGGCGACCGGCACTGATCTGCCAACCACTACACCACTACCGGAGC	540
DB	481 ATGGCCCTGAAGGTGGGCGACCGGCACTGATCTGCCAACCACTACACCACTACCGGAGC	540
QY	541 AAGAGGCCGTGCGCGCCCTGACATGCCCCGCTTCCACTTCAACCGACATCCGGCTCCAG	600
DB	541 AAGAGGCCGTGCGCGCCCTGACATGCCCCGCTTCCACTTCAACCGACATCCGGCTCCAG	600
QY	601 ATGCTGCGGAGAGAAAGGACGAGTACTTCAAGCTGTACGAGGCCAGCGTGGCCGGTAC	660
DB	601 ATGCTGCGGAGAGAAAGGACGAGTACTTCAAGCTGTACGAGGCCAGCGTGGCCGGTAC	660
QY	661 AGCGACCTGCCCCGAGAAAGGCCAATGA 687	
DB	661 AGCGACCTGCCCCGAGAAAGGCCAATGA 687	
RESULT 2		
LOCUS	AX527900	687 bp DNA linear PAT 21-NOV-2002
DEFINITION	Sequence 13 from Patent WO0230965.	
ACCESSION	AX527900	
VERSION	AX527900.1	GI:25172349
KEYWORDS		
SOURCE	Heteractis crispa (leathery sea anemone)	
ORGANISM	Heteractis crispa	
	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;	
	Stichodactylidae; Heteractis.	
REFERENCE	1	

AUTHORS		Lukyanov, S.A., Fradkov, A.F., Lukyanov, K.A. and Gurskaya, N.G.									
TITLE		Nucleic acids encoding stichodactylidae chromoproteins									
JOURNAL		Patent: WO 0230965-A 13 18-Apr-2002;									
FEATURES		Clontech Laboratories Inc. (US)									
source		Location/Qualifiers									
		1. .687									
		/organism="Heteractis crispa"									
		/mol_type="unassigned DNA"									
		/db_xref="taxon:175771"									
ORIGIN											
Query Match		97.7%; Score 671; DB 6; Length 687;									
Best Local Similarity		98.5%; Pred. No. 2.5e-77;									
Matches 677; Conservative		0; Mismatches 10; Indels 0; Gaps 0;									
QY	1	ATGTGTAGCGCGCTGTGAAGAGAGATATGCGCATCAAGATGTACATGAGGGGACCGTG	60								
DB	1	ATGTGTAGCGCGCTGTGAAGAGAGAGATGCGCATCAAGATGTACATGAGGGGACCGTG	60								
QY	61	AACGECCTACTTCTCAAGTGGGAGGGGCGAGGGCGACGGCAACCCCTTCCGCCGGCACCCAG	120								
DB	61	AACGGCTACTACTTCAAGTGGGAGGGGCGAGGGCGACGGCAACCCCTTCCGCCGGCACCCAG	120								
QY	121	AGCATGAGAATCCACGTGACCGAGGGCGCCCCCTTGCCCTTGGCTTGCATCCTGGCC	180								
DB	121	AGCATGCGGATCCACGTGACCGAGGGCGCCCCCTTGCCCTTGGCTTGCATCCTGGCC	180								
QY	181	CCCTGCTGCGAGTACGGGACAGGACCTTGTGCAACACACCGCCGAGATCCCGACTTC	240								
DB	181	CCCTGCTGCGAGTACGGGACAGGACCTTGTGCAACACACCGCCGAGATCCCGACTTC	240								
QY	241	TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGGAGAGAACAACCACTACGAGGACGGC	300								
DB	241	TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGGAGAGAACAACCACTACGAGGACGGC	300								
QY	301	GGCATCTGACCGCCACCAAGACACAGCCTGGAAGGGCAACTGCTGATCTACAAGGTG	360								
DB	301	GGCATCTGACCGCCACCAAGACACAGCCTGGAAGGGCAACTGCTGATCTACAAGGTG	360								
QY	361	AAGGTGACGGGACCAACTTTCGCCGACGGCCCCCTGATGAAGAACAGAGCGCGGC	420								
DB	361	AAGGTGCTGGGCAACCACTTTCGCCGACGGCCCCCTGATGAAGAACAGAGCGCGGC	420								
QY	421	TGGAGGCCACGACCGGAGGTGTGTACCCCGAGAAACGGCGTGTGCGGCCGGAACGTG	480								
DB	421	TGGAGGCCACGACCGGAGGTGTGTACCCCGAGAAACGGCGTGTGCGGCCGGAACGTG	480								
QY	481	ATGGCCCTGAAGGTGGGCGACCGGCACTGATCTGCCACCACTACACCACTACCGGAGC	540								
DB	481	ATGGCCCTGAAGGTGGGCGACCGGCGCTGATCTGCCACCACTACACCACTACCGGAGC	540								
QY	541	AAGAAGCGCTGCGCGCCCTGACCATGCCCCGCTTCCACTTACCGACATCCGGCTCCAG	600								
DB	541	AAGAAGCGCTGCGCGCCCTGACCATGCCCCGCTTCCACTTACCGACATCCGGCTCCAG	600								
QY	601	ATGCTGCGGAAGAAGAAGACGAGTACTTTCAGCTGTACGAGGCCAGCGTGGCCCCGTAC	660								
DB	601	ATGCTGCGGAAGAAGAAGACGAGTACTTTCAGCTGTACGAGGCCAGCGTGGCCCCGTAC	660								
QY	661	AGCGACTGCCCCGAGAGAGGCCCACTGA	687								
DB	661	AGCGACTGCCCCGAGAGAGGCCCACTGA	687								
RESULT 3											
LOCUS		AX527902 1396 bp DNA linear PAT 21-NOV-2002									
DEFINITION		Sequence 15 from Patent WO0230965.									
ACCESSION		AX527902									
VERSION		AX527902.1 GI:25172350									
KEYWORDS		synthetic construct									
SOURCE		synthetic construct									
ORGANISM		synthetic construct									

REFERENCE	1	artificial sequences.
AUTHORS	Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.	
TITLE	Nucleic acids encoding stichodactylidae chromoproteins	
JOURNAL	Patent: WO 0230965-A 15 18-APR-2002;	
Clontech Laboratories Inc. (US)		
FEATURES	location/Qualifiers	
source	1..1396	
	/organism="synthetic construct"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:32630"	
	/note="fusion construct"	
ORIGIN		
Query Match	97.7%; Score 671; DB 6; Length 1396;	
Best Local Similarity	98.5%; Pred. No. 2.2e-77;	
Matches 677; Conservative	0; Mismatches 10; Indels 0; Gaps 0;	
QY	1 ATGCTGAGCGGCTGCTGTAAGAGAGTATGCGCATCAAGATGTACATGAGGGCAACCGTG	60
Db	710 ATGCTGAGCGGCTGCTGTAAGAGAGAGCATGCGCATCAAGATGTACATGAGGGCAACCGTG	769
QY	61 AACGGCCACTACTTCAAGTGGCGAGGGCGAGCGGCAACCCCTTGCCGGCAACCCAG	120
Db	770 AACGGCCACTACTTCAAGTGGCGAGGGCGAGCGGCAACCCCTTGCCGGCAACCCAG	829
QY	121 AGCATGAGAATCCACGTGACCGAGGGCGCCCTGCCCCCTTGCCCTTGACATCCTGGCC	180
Db	830 AGCATGCGGATCCACGTGACCGAGGGCGCCCTGCCCCCTTGCCCTTGACATCCTGGCC	889
QY	181 CCCTGCTCGAGTACCGGACAGGACCTTGTCACCAACCGCCGAGATCCCGACTTC	240
Db	890 CCCTGCTCGAGTACCGGACAGGACCTTGTCACCAACCGCCGAGATCCCGACTTC	949
QY	241 TTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCACTTACGAGGACGGC	300
Db	950 TTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCACTTACGAGGACGGC	1009
QY	301 GGCATCTTGACCGCCCAACGAGACACCAAGCCTGAGGGCAACTGCTGATCTACAAGTG	360
Db	1010 GGCATCTTGACCGCCCAACGAGACACCAAGCCTGAGGGCAACTGCTGATCTACAAGTG	1069
QY	361 AAGGTGACGGCAACCACTTCCCCCGGAGCGGCCCGTGATGAAGAACAAAGCGGCGGC	420
Db	1070 AAGGTGCTGGGCAACCACTTCCCCCGGAGCGGCCCGTGATGAAGAACAAAGCGGCGGC	1129
QY	421 TGGAGCCGACGACCGAGGTGTATCCCGAGAAACGCGCTGTGCGGCCGGAACGTG	480
Db	1130 TGGAGCCGACGACCGAGGTGTATCCCGAGAAACGCGCTGTGCGGCCGGAACGTG	1189
QY	481 ATGGCCCTGAAGGTGGGGCGACCGGCACTGATCTGCCAACCACTACCAAGTACCAGAGC	540
Db	1190 ATGGCCCTGAAGGTGGGGCGACCGGCGGCTGATCTGCCAACCACTACCAAGTACCAGAGC	1249
QY	541 AAGAAGGCGGTGCGGCCCTTGACCATGCGCGGCTTCCACTTCAACGACATCCGGCTCCAG	600
Db	1250 AAGAAGGCGGTGCGGCCCTTGACCATGCGCGGCTTCCACTTCAACGACATCCGGCTCCAG	1309
QY	601 ATGCTGCGGAAGAAGAGCAGTACTTGAGCTGTACGAGGCCAGCTGGCCCGGTAC	660
Db	1310 ATGCTGCGGAAGAAGAGCAGTACTTGAGCTGTACGAGGCCAGCTGGCCCGGTAC	1369
QY	661 AGCGACTGCCCCGAGAAGGCCAAGTGA	687
Db	1370 AGCGACTGCCCCGAGAAGGCCAAGTGA	1396
RESULT 4		
AX527904		
LOCUS	AX527904	1424 bp DNA linear PAT 21-NOV-2002
DEFINITION	Sequence 17 from Patent WO0230965.	
AX527904		
AX527904.1	GI:25172351	

KEYWORDS	synthetic construct	
SOURCE	synthetic construct	
ORGANISM	artificial sequences.	
REFERENCE	1	
AUTHORS	Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.	
TITLE	Nucleic acids encoding stichodactylidae chromoproteins	
JOURNAL	Patent: WO 0230965-A 17 18-APR-2002;	
Clontech Laboratories Inc. (US)		
FEATURES	location/Qualifiers	
source	1..1424	
	/organism="synthetic construct"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:32630"	
	/note="fusion construct"	
ORIGIN		
Query Match	97.4%; Score 669.4; DB 6; Length 1424;	
Best Local Similarity	98.4%; Pred. No. 3.5e-77;	
Matches 676; Conservative	0; Mismatches 11; Indels 0; Gaps 0;	
QY	1 ATGCTGAGCGGCTGCTGAAGAGAGTATGCGCATCAAGATGTACATGAGGGCAACCGTG	60
Db	14 ATGCTGAGCGGCTGCTGAAGAGAGAGCATGCGCATCAAGATGTACATGAGGGCAACCGTG	73
QY	61 AACGGCCACTACTTCAAGTGGCGAGGGCGAGCGGCAACCCCTTGCCGGCAACCCAG	120
Db	74 AACGGCCACTACTTCAAGTGGCGAGGGCGAGCGGCAACCCCTTGCCGGCAACCCAG	133
QY	121 AGCATGAGAATCCACGTGACCGAGGGCGCCCTTGCCCTTGCCCTTGACATCCTGGCC	180
Db	134 AGCATGCGGATCCACGTGACCGAGGGCGCCCTTGCCCTTGCCCTTGACATCCTGGCC	193
QY	181 CCCTGCTCGAGTACCGGACAGGACCTTGTCACCAACCGCCGAGATCCCGACTTC	240
Db	194 CCCTGCTCGAGTACCGGACAGGACCTTGTCACCAACCGCCGAGATCCCGACTTC	253
QY	241 TTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCACTTACGAGGACGGC	300
Db	254 TTCAAGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCACTTACGAGGACGGC	313
QY	301 GGCATCTTGACCGCCCAACGAGACACCAAGCCTGAGGGCACTGCTGATCTACAAGTG	360
Db	314 GGCATCTTGACCGCCCAACGAGACACCAAGCCTGAGGGCACTGCTGATCTACAAGTG	373
QY	361 AAGGTGACGGCAACCACTTCCCCCGGAGCGGCCCGTGATGAAGAACAAAGCGGCGGC	420
Db	374 AAGGTGCTGGGCAACCACTTCCCCCGGAGCGGCCCGTGATGAAGAACAAAGCGGCGGC	433
QY	421 TGGAGCCGACGACCGAGGTGTATCCCGAGAAACGCGCTGTGCGGCCGGAACGTG	480
Db	434 TGGAGCCGACGACCGAGGTGTATCCCGAGAAACGCGCTGTGCGGCCGGAACGTG	493
QY	481 ATGGCCCTGAAGGTGGGGCGACCGGCACTGATCTGCCAACCACTACCAAGTACCAGAGC	540
Db	494 ATGGCCCTGAAGGTGGGGCGACCGGCGGCTGATCTGCCAACCACTACCAAGTACCAGAGC	553
QY	541 AAGAAGGCGGTGCGGCCCTTGACCATGCGCGGCTTCCACTTCAACGACATCCGGCTCCAG	600
Db	554 AAGAAGGCGGTGCGGCCCTTGACCATGCGCGGCTTCCACTTCAACGACATCCGGCTCCAG	613
QY	601 ATGCTGCGGAAGAAGAGCAGTACTTGAGCTGTACGAGGCCAGCGTGCCCGGTAC	660
Db	614 ATGCTGCGGAAGAAGAGCAGTACTTGAGCTGTACGAGGCCAGCGTGCCCGGTAC	673
QY	661 AGCGACTGCCCCGAGAAGGCCAAGTGA	687
Db	674 AGCGACTGCCCCGAGAAGGCCAAGTGA	700
RESULT 5		
AX527894		
LOCUS	AX527894	681 bp DNA linear PAT 21-NOV-2002

DEFINITION	Sequence 7 from Patent WO0230965.
ACCESSION	AX527894
VERSION	AX527894.1
KEYWORDS	GI:25172346
SOURCE	.
ORGANISM	Heteractis crispata (leathery sea anemone)
	Heteractis crispata
	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria; Stichodactylidae; Heteractis.
REFERENCE	1
AUTHORS	Lukyanov, S.A., Pradkov, A.F., Lukyanov, K.A. and Gurskaya, N.G.
TITLE	Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL	Patent: WO 0230965-A 7 18-APR-2002; Clontech Laboratories Inc. (US)
FEATURES	Location/Qualifiers
	1..681
source	/organism="Heteractis crispata"
	/mol_type="unassigned DNA"
	/db_xref="taxon:175771"
ORIGIN	

Query Match	64.2%;	Score 441.2;	DB 6;	Length 681;
Best Local Similarity	78.2%;	Pred. No. 8.9e-48;		
Matches 530;	Conservative 0;	Mismatches 148;	Indels 0;	Gaps 0
QY	10	GGCCTGCTGAAGAGAGATATGCGCATCAAGATGTACATGGAGGGCAACCGTGAAACGGCCAC	69	
Db	4	GGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAGGGCACGGTTAATGGCCAT	63	
QY	70	TACTTCAAGTGCAGGGCGGAGGGCGACGGCAACCCCTTCGCCGCAACCCAGACATGAGA	129	
Db	64	TATTTCAAGTGTGAAGAGAGAGGGAGACGGCAACCCATTTGCAGGTACGACAGCATGAGG	123	
QY	130	ATCCACGTGACCGAGGGGCGCCCCCTTGCCCTTGCCCTTCGACATCCTGGCCCCCTGCTGC	189	
Db	124	ATTATGTACACCGAAGGGGCTCCATTCACATTTGCTTCGACATTTGGCACCGTGTGT	183	
QY	190	GAGTACGGCAGCAGGAGACCTTCGTGACACACACCGCCGAGATCCCGCACTTCTTCAAGCAG	249	
Db	184	GAGTACGGCAGCAGGAGACCTTTGTCCACCATACGGCAGAGATTCGGAATTTCTTCAAGCAG	243	
QY	250	AGCTTCCCCGAGGCTTACCTGGGAGAGAAACCAACCTACGAGACGGCGGCAATCCTG	309	
Db	244	TCTTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAGATGAGGCAATCTT	303	
QY	310	ACCGCCCAACGAGACACAGAGCTGAGAGGGCAACTGCTGATCTACAGGTGAAAGGTGCAC	369	
Db	304	ACTGCTCATCAGACACAAAGCCTGAGAGGGAACTGCTTATATACAAAGTGAAAGTCTCTT	363	
QY	370	GGCACCAACTTCCCCCGCCGACGGCCCCGTGATGAGAAACAAGAGCGCGGTGGAGCCCC	429	
Db	364	GGTACCAATTTCTCTGCTGATGGCCCCGTGATGAGAAACAATACAGAGATGGAGCCA	423	
QY	430	AGCACCGAGGTGTGTACCCCGAAGACGGCGTGTGTGCGGCCGGAACGTGATGGCCCTG	489	
Db	424	AGCACTGAGGTGTTTATCCAGAGAAATGTTGTCCTGTGTGACCTAATGTGATGACCCTT	483	
QY	490	AAGGTGGGCGACCGGCACCTGATCTGCCACACACTACACCAAGCTACCGGAGCAAGAGCC	549	
Db	484	AAAGTCGGTGATCGTCTGTTGATCTGCCATCACTATACTTCTTACAGGTCCAGAAAGCA	543	
QY	550	GTGCGGCGCTGATACCATGCCCCGCTTCCACTTCAACGACATCCGGCTCCAGATGCTGCGG	609	
Db	544	GTCCGTGCTTGAACAATGCCAGGATTTCAATTTACAGACATCCGCTTCAGATGCTGAGG	603	
QY	610	AAGAAAGAGCAGAGTACTTTCGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGGCACTG	669	
Db	604	AAAGAGAAAGCAGAGTACTTTTGAACGTACGAGCATCTGTGGCTAGTACAGTGAATCTT	663	
QY	670	CCCGAGAGAGGCCAACTGA 687		
Db	664	CCTGAAAAAGCAATTGA 681		

RESULT	6
AX527896	
LOCUS	
DEFINITION	AX527896 681 bp DNA linear PAT 21-NOV-2002
ACCESSION	Sequence 9 from Patent WO0230965.
VERSION	AX527896
KEYWORDS	AX527896.1 GI:25172347
SOURCE	.
ORGANISM	Heteractis crispa (leathery sea anemone)
	Heteractis crispa
	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
	Stichodactylidae; Heteractis.
REFERENCE	1
AUTHORS	Lukyanchov,S.A., Fradkov,A.F., Lukyanchov,K.A. and Gurskaya,N.G.
TITLE	Nucleic acids encoding stichodactylidae chromoproteins
JOURNAL	Patent: WO 0230965-A 9 18-APR-2002;
	Clontech Laboratories Inc. (US)
FEATURES	Location/Qualifiers
source	1..681
	/organism="Heteractis crispa"
	/mol_type="unassigned DNA"
	/db_xref="taxon:175771"
ORIGIN	

Query Match	64.0%;	Score 439.6;	DB 6;	Length 681;
Best Local Similarity	78.0%;	Pred. No. 1.4e-47;		
Matches 529;	Conservative 0;	Mismatches 149;	Indels 0;	Gaps 0;
QY	10	GGCCTGCTGAAGAGAGATATGCGCATCAAGATGTACATGAGGGGCAACCGTGAACGGCCAC	69	
Db	4	GCTTTGTTGAAGAAGATATGCGCATCAAGATGTACATGGAAGGCAACGGTTAATGGCCAT	63	
QY	70	TACTTCAAGTCGAGGGGCGGAGCGACGCAACCCCTTTCGCCGCAACCCAGAGCATGAGA	129	
Db	64	TATTTCAAGTGTGAAGAGAGGGAGACGGCAACCCATTGTGCAGGTACGCAGAGCATGAGG	123	
QY	130	ATCCACGTGACCGAGGGGCGCCCCCTGCCCCCTTGCCTTGCATCTTGCCCCCTGCTGC	189	
Db	124	ATTCATGTCAACGAAAGGGGCTTCATTACCATTTGCTTGCACATTTTGGCACCCGTGTGT	183	
QY	190	GAGTACGGGACAGGACCTTTCGTGCACCAACCGCCGAGATCCCGACTTCTTCAAGCAG	249	
Db	184	GCGTACGGGACAGGACCTTTCGTGCACCAATACGGCAGAGATCCCGATTCTTCAAGCAG	243	
QY	250	AGCTTCCCGCAGGGCTTCACCTGGGAGAGAACCAACCACTTACGAGGACGGCGCATCCTG	309	
Db	244	TCTTTCCTGTAAGGCTTTACTTTGGAAAGAACCAACCTTATGAAGATGAGGCACTTCTT	303	
QY	310	ACCGCCACACGAGACACCAAGCCTGGAGGGCACTGCCGTGATCTACAAGGTGAAGTGCAC	369	
Db	304	ACTGCTCATACGAGACACCAAGCCTGGAGGGGAACTGCCCTTATATCAAGGTGAAGTCCCTT	363	
QY	370	GGCACCAACTCCCCCGCCGACGGCCCCGTGATGAAGAACAAGCGGGCGGTGGAGCCC	429	
Db	364	GGTACCAATTTCTGCTGATGGCCCCCGTGAAGAACAATCAAGAGGATGGAGCCA	423	
QY	430	AGCACCGAGTGTGTACCCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGATGGCCCTG	489	
Db	424	AGCACTGAGTGTGTATTCAGAGAATGTGTCTGTGTGACGTATGTGATGGCCCTT	483	
QY	490	AAGTGGGGCACCGGCACTGATCTGCCAACCACTACACCAAGCTAACCGGAGCAAGAAGCC	549	
Db	484	AAAGTCGGTATCGTCTGTGATCTGCCAATCACTATCTTTACAGGTCCAGAAAGCA	543	
QY	550	GTGCGCGCCCTGACCATGCCCGGCTTCCACTTCAACGACATCCGGCTCCAGATGCTGGG	609	
Db	544	GTCGCTGCTTGACAATGCCAGGATTTCAATTTACAGACATCCGCCCTTCAGATGCTGAGG	603	
QY	610	AAGAAGAGAGACGAGTACTTGCAGCTGTACGAGGCGCAGCGTGGCCCGGTACAGCAGCTG	669	
Db	604	AAAGAGAAAGACGAGTACTTGAAGTGTACGAAGCATCTGTGGCTAGGTACAGTATCTT	663	
QY	670	CCCGAAGAGGCCCACTGA	687	



Db	664	CCTGAAAAAGCAATTGA	681
RESULT 7			
AX527910			
LOCUS	AX527910	684 bp	DNA
DEFINITION	Sequence 23 from Patent WO0230965.		linear
ACCESSION	AX527910		PAT 21-NOV-2002
VERSION	AX527910.1	GI:25172356	
KEYWORDS			
SOURCE			
ORGANISM	Heteractis crispa (leathery sea anemone)		
	Heteractis crispa		
	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;		
	Stichodactylidae; Heteractis.		
REFERENCE			
AUTHORS	1 Lukanov,S.A., Pradkov,A.F., Lukanov,K.A. and Gurskaya,N.G.		
TITLE	Nucleic acids encoding stichodactylidae chromoproteins		
JOURNAL	Patent: WO 0230965-A 23 18-APR-2002;		
	Clontech Laboratories Inc. (US)		
FEATURES			
source	1..684		
	location/Qualifiers		
	/organism="Heteractis crispa"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:175771"		
ORIGIN			
Query Match	63.8%;	Score 438.2;	DB 6; Length 684;
Best Local Similarity	77.6%;	Pred. No. 2.2e-47;	
Matches	530; Conservative	0; Mismatches 153; Indels	0; Gaps 0;
QY	5	TGAGCGGCGCTGTGAAGAGAGATGCGCATCAAGATGTACATGGAAGGCAACCGTGAACG	64
Db	2	TGCGTGGTTGTGTGAAGAAGATATGCGCATCAAGATGTACATGGAAGGCAACCGTGAATG	61
QY	65	GCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCA	124
Db	62	GCCATTATTTCAAAGTGTGAAGAGAGAGGAGACGGCAACCCATTTCAGGTACGACAGCA	121
QY	125	TGAGAAATCCACGTGACCGAGGGCGCCCCCTGCCCCCTTGCCCTTGACATCTGCCCCCT	184
Db	122	TGAGGATTGATGTCACCGAAGGGGCTCCATTACCATTTGCCCTTGACATTTTGACACCGT	181
QY	185	GCTGCGAGTACGGCAGCAGACCTTCGTGACACCACACCGCGGAGATCCCCGACTTCTCA	244
Db	182	GTTGTGAGTACGGCAGCAGACCTTGTGCCACCATACGGCAGAGATTCCCGATTCTTCA	241
QY	245	AGCAGAAGCTTCCCCGAGGGCTTCACTGGGAGAGAACCAACCACTACGAGGACGGCGCA	304
Db	242	AGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAAGATGAGGCA	301
QY	305	TCCTGACCGCCCAACGACACAGCCTGAGGGCAACTGCTGATCTTACAAGGTGAAG	364
Db	302	TTCTTACTGCTCATCAGACACAAGCTGAGGGGAACTGCTTATATACAAAGTGAAG	361
QY	365	TGCACGGCACCAACTTCCCCCGGACGGCCCCCGTGATGAAGAACAGCGCGGCTGG	424
Db	362	TCCTTGGTACCAATTTTCTGCTGATGGCCCCGTGATGAAGAACAAATCAGAGGATGG	421
QY	425	AGCCACGACCGAGGTGTATCCCCGAGAACGCGCTGTGGCGCCGGAACGTGATGG	484
Db	422	AGCCACGACCTGAGGTGTATCCAGAGATGTGTCTCTGTGAGACGTATGTGATGG	481
QY	485	CCCTGAAGGTGGGCGACCGGCACTGATCTGCCACCACTTACCAAGCTAACCGGACAGA	544
Db	482	CCCTTAAAGTCGGTGATCGTGTGATCTGCCATCTCTATACTTCTTACAGGTCCAAGA	541
QY	545	AGGCGGTGCGGCGCCCTGACCATGCGCGGCTTCCACTTACCGCATCCGGCTCCAGATGC	604
Db	542	AAGCAGTCCGTGCTTGAACAATGCCAGGATTTCAATTTACAGCATCCGCTTCAGATGC	601
QY	605	TGCGGAAGAAGACGAGTACTTCAGCTGTACGAGGCCAGCGTGCCCGGTACAGCG	664
Db	602	CGAGAAAAAGAAAGACGAGTCTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG	661

[illegible]

[illegible][illegible]

REFERENCE	1 (bases 1 to 908)
AUTHORS	Gurskaya, N. G., Fradkov, A. F., Tersikh, A., Matz, M. V., Labas, Y. A., Martynov, V. I., Yanushovich, Y. G., Lukyanov, K. A. and Lukyanov, S. A.
TITLE	GFP-like chromoproteins as a source of far-red fluorescent proteins
JOURNAL	FEBS Lett. 507 (1), 16-20 (2001)
MEDLINE	21538626
PUBMED	11682051
REFERENCE	2 (bases 1 to 908)
AUTHORS	Gurskaya, N. G., Lukyanov, K. A., Labas, Y. A. and Lukyanov, S. A.
TITLE	Direct Submission
JOURNAL	Submitted (23-MAR-2001) Institute of Bioorganic Chemistry, Miklukho-Maklaya 16/10, Moscow 117997, Russia
FEATURES	

```

FEATURES
  source      location/Qualifiers
              1..908
              /organism="Heteractis crispa"
              /mol_type="mRNA"
              /db_xref="taxon:175771"
              77..760
CDS

```

```

/note="hccp"
/codon_start=1
/product="GFP-like chromoprotein"
/protein_id="AAL27538.1"
/db_xref="GI:16660116"
/translation="MAGLLKESMRIKMYMEGVANGHYFKCEGEGDGNPFTGTQSMRIH
VTEGAPLPFAFDILAPCCCEYSRVEVHTAEIDPFKQSFPEPFTWERTTYEDGKITL
TAHQDTSLEGNCILYKVKVLGNTNPADGPMKKSQGWPECTEVVYPENGVLGGRNVM
ALKVQDRRLICHLVTSYRSKKAVALTMPGFHTDIRLQMPRKKKDEYFELYEASVAR
YSDLPBEKAN"

```

**ORIGIN**

Query Match	63.6%	Score 436.6;	DB 3;	Length 908;
Best Local Similarity	77.5%	Pred. No. 3.3e-47;		
Matches 529; Conservative	0;	Mismatches 154;	Indels 0;	Gaps 0;

Dy  
5 TGAGCGGCTGTGAAGAGATATGCCCATCAAGTGTACTGAGGGCACCCTGAACG 64

Dδ  
78 TGGCTGTTTCTTGAAAGAATAATGCCCATCAAGATGTAATGGAAGGCACCGTTAATG 137

QY      65 GCCACTACTTCAAGTGCCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCA 124  
|||||  
|||  
Db    138 GCCATTATTTCAAGTGTGAAGAGAGAGGAGACGGCAACCATTTACAGGTACGCAGAGCA 197

**QY** 125 TGAGATCCACGTGACCGAGGGCGCCCCCTGCCTTCGACATCCTGGCCCCCT 184  
|||||  
**D6** 198 TGAGGATTCATGTCAACCGAGGGGCTCCATTACCATTGGCTTCCGATTCTTTGGGCT 257  
|||||

[illegible]

**QY** 245 AGCAGAGCTTCCCCGAGGGCTTCACCTTGGAAGAACCACCACTACTGAGAGCGCGCA 304  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
**bB** 318 AGCATCTTTCCCTGCCTATCTGCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTA 317  
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY	305	TCCTGACCCGCCACCAGAGACACCCAGCCTGGAGGGCAACTGCCTGATCTACAAAGTGAAG	364
Db	378	TTCTTACTGCTCATCAGGACACAAGCCTGGAGGGAACTGCCTTATATACAAAGTGAAG	437
QY	365	TGCACGGGACCAACTTCCCGCCGCCAGCGCCCCCGTGATGAAGAACAAGACCGCGGCTGG	424
Db	438	TCCTTGGTACCAATTTTCCTGCTGATGGCCCCCGTGATGAAGAACAAATCAGGAGGATGG	497
QY	425	AGCCACGACACCGAGGTGTGTACCCCGAGAAACGCGCTGTGTGCGGCCGAACGTGATGG	484
Db	498	AGCCATGCACCTGAGGTGGTTTATTCAGAGAAATGTGTCTGTGTGACGTAATGTGATGG	557
QY	485	CCCTGAAGGTGGGCGCACCGGACCTGATCTGCCCACTACACCACTACCCGAGACAAGA	544
Db	558	CCCTTAAAGTCGGTGATCGTCGTTGATCTGCCACTCTATATCTTACAGGTCACAAGA	617
QY	545	AGGCGGTGCGCGCCCTGACCATGCCCCGGCTTCCACTTACCGGACATCCGGCTCCAGATGC	604
Db	618	AAGCAGTCCGTGCTTGACAATGCGAGGATTTCAATTTACAGACATCCGCTTCAGATGC	677
QY	605	TGCGGAAGAAGAAGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGCCCCGGTACAGCG	664
Db	678	CGAGGAAAAAAGAAAGACGAGTACTTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG	737
QY	665	ACTGCGCCGAGAAGGCCCACTGA	687
Db	738	ATCTTCTCTGA AAAAGCAAAATTGA	760

RESULT 10			
AX527888			
LOCUS	AX527888	910 bp	DNA
DEFINITION	Sequence 1 from Patent WO0230965.	linear	PAT 21-NOV-2002

ACCESSION	AX527888
VERSION	AX527888.1
KEYWORDS	GI:25172343
SOURCE	.
	<i>Heteractis crispa</i> (leathery sea anemone)

ORGANISM	REFERENCE
<i>Heteractis crispata</i>	Lukyanov, S. A., Fradkov, A. F., Lukyanov, K. A. and Gurskaya, N. G.
Eukaryota; Metazoa; Chnidaria; Anthozoa; Zoantharia; Actiniaria;	Nucleic acids encoding stichodactylid chromoproteins
Stichodactylidae; Heteractis.	Patent: WO 0230965-A 1 18-APR-2002;
1	Clontech Laboratories Inc. (USA)

FEATURES	source
Location/Qualifiers	1. .910

```

/organism="Heteractis crispa"
/mol_type="unassigned DNA"
/db_xref="taxon:175771"

```

Query Match	63.3%;	Score 435;	DB 6;	Length 910;
Best Local Similarity	77.3%;	Pred. No. 5.3e-47;		
Matches 528;	Conservative	0;	Mismatches 155;	Indels 0;
			Gaps	0;

OY            5 TGAAGCGCCTGCTGAAGAGAGATGCCCATCAACATGTACATGGAGGGCACCGTTGAACG 64  
             || | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db            83 TGCGTGGTTTGTTGAAGAAGAAAGTAGTCGCATCAGATGTACATGGAGGGCACGGTTAATG 142

QY 65 GCCACTACTTCAAGTCGAGGCGGAGGGCGACGGCAACCCCTTGGCCGACCCAGACA 124  
 |||||  
 Ddb 143 GCCATTATTTCAAGTGTGAAGGAGAGGAGACCGCAACCCATTACAGGTATCCGAGAGCA 202  
 |||||

QY 125 TGAGATCCACGTGACCGAGGGCCCCCCTGCGCTTCGACATCCTGGCCCCCT 184  
 |||||  
 Db 203 TGAGGATTCATGACCGAAGGGGCTCCATTACCAATTTGGCTTTGGCAATTTTGGTACCCCT 263

185 GCTGCGAGTACGGCAGCAGA CTTTCGTGCACCACACCGCCGAGATCCCCGACTTCTTCA 244  
| | | | | | | | | | | | | | | | | | | | | |  
263 GTTGTGAGTACGGCAGCAGA CCTTTCGTGCACCACACCGCCGAGATCCCCGACTTCTTCA 260

QY	245	AGCAGAGCTTCCCCGAGGGCTTCACTGGGAGAGAAACCACTACGAGGACGGCGCA	304
Db	323	AGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAAGAACCAACTATGAAGATGAGGCA	382
QY	305	TCCTGACCGCCCAACGAGACACAGCTGGAGGGCACTGCTATCTACAAGTGAAAG	364
Db	383	TTCTTACTGCTCATCAGACACAAGCTGGAGGGAACTGCTTATATACAAGGTGAAG	442
QY	365	TGCACGGCACCAACTTCCCCCGCAGCGCCCCGTATGAAGACAAGCGCGCTGGG	424
Db	443	TCCTTGGTACCAATTTTCTGTGATGGCCCCGTATGAAGACAATCAGAGATGGG	502
QY	425	AGCCAGCACCGAGGTGTGTACCCCGAGAACGCGTGTGCGGCCGAACTGATGG	484
Db	503	AGCCATGCACGTAGGTGTATCCAGAAGATGGTGTCTGTGTGAGACGTAATGTGATGG	562
QY	485	CCCTGAAGGTGGGGAACCGGCACTGATCTGCCACCACTACAGCTAACGGAGCAAGA	544
Db	563	CCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCTCTATCTTCTTACAGSTCCAAGA	622
QY	545	AGGCGGTGGCGCCCTGACCATGCCCCGGCTTCCACTTCAACGACATCCGGCTCCAGATGC	604
Db	623	AAGCAGTCCGTGCTTGAACAATGCCAGGATTTCTATTTTACAGACATCCGCCCTTCAGATGC	682
QY	605	TGCGGAAGAAGAAGGACGAGTACTTGAAGCTGTACGAGGCCAGCGTGGCCCCGATACAGCG	664
Db	683	CGAGGAAAAAGAAAGACGAGTACTTTGAACCTGTACGAAGCATCTGTGGCTAGGTACAGTG	742
QY	665	ACCTGCCCGAGAGGCCCACTGA	687
Db	743	ATCTTCTGAAAAAGCAATTGA	765

RESULT	11			
AX527914				
LOCUS	AX527914	910 bp	DNA	linear
DEFINITION	Sequence 27 from Patent WO0230965.			PAT 21-NOV-2002
ACCESSION	AX527914			
VERSION	AX527914.1	GI:25172358		
KEYWORDS	.			
SOURCE	Heteractis crispa (leathery sea anemone)			
ORGANISM	Heteractis crispa			
	Bukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria; Stichodactylidae; Heteractis.			
REFERENCE	1			
AUTHORS	Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.			
TITLE	Nucleic acids encoding stichodactylidae chromoproteins			
JOURNAL	Patent: WO 0230965-A 27 18-APR-2002;			
	Clontech Laboratories Inc. (US)			
FEATURES	Location/Qualifiers			
source	1..910			
	/organism="Heteractis crispa"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:175771"			

**ORIGIN**

Query Match	63.3%	Score 435;	DB 6;	Length 910;
Best Local Similarity	77.3%	Pred. No. 5.3e-47;		
Matches 528; Conservative	0;	Mismatches 155;	Indels 0;	Gaps 0;

QY	5	TCAGCGGCCCTGCTGAAGGAGAGATGCCGATCAAGATGTACATGGAGGGCACCGTGAACG	64
Db	83	TGGCTGGSTTTGTTGAAAGAAAGTATGCCGATCAGATGTACATGGAAAGGCACGGTTAATG	142
QY	65	GCCCACTACTTCAAGTGCAGGCGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCA	124
Db	143	GCCATTATTTCAAGTGTGAAGGAGAGAGGAGACGGCAACCCATTACAGGTACCGCAGAGCA	202
QY	125	TGAGAAATCCACGTCGACCCGAGGGCGCCCCCTGCGCTTCGACATCCTGGCCCCCT	184
Db	203	TGAGGATTCATGTCTACCCGAAAGGGGCTCCATTACCAATTGCGCTTCGACATTTGGCACCGT	262
QY	185	GCTGCGAGTACGGCGACGAGACCTTCGTGACCAACCGCGAGATCCCCGACTTCTTCA	244

Db	263	GTGTGAGTACGGCAGCAGGACCTTTGTCCACCATTACGGCAGAGATTCCCGATTCTTCA	322
QY	245	AGCAGAGCTTCCCCGAGGGCTTCACCTGGAGAGAAACCACCTACGAGGACGGCGCA	304
Db	323	AGCAGTCTTCCCTGGAAGGCTTTACTTGGAAAGAACCAACCACTTATGAAGATGAGGCA	382
QY	305	TCCTGACCGCCACCAAGACAACGACCTGGAGGGCAATGCTGATCTACAAGGTGAAG	364
Db	383	TTCTTACTGCTCATCAGACACAAGCCTGAGAGGGAATGCTTATATACAAGGTGAAG	442
QY	365	TGCACGGCACCAACTTCCCCGCGCAGCGCCCGGTGATGAAGAACAGAGCGCGCTGGG	424
Db	443	TCCTTGGTACCAATTTCTGTGATGGCCCCGTGATGAAGAACAAATCAGAGGATGGG	502
QY	425	AGCCCAAGCACCAGGTGTGTATCCCCGAGAACGGCGTGTGCGCCGGAACGTGATGG	484
Db	503	AGCCATGCACGTGAGGTGTTTATCCAGAGAAATGTTCTGTGTGACGTAAATGTGATGG	562
QY	485	CCCTGAAGGTGGGCGACCGGCACCTGATCTGCCACCACTACACAGCTACCGGACAAGA	544
Db	563	CCCTTAAAGTCGGTGAATCGTGTGATCTGCCATCTTAACTTCTTACAGGTCCAAGA	622
QY	545	AGCGCGTGCAGCGCCCTGACCATGCCCCGCTTCCAATTACCGACATCCGGCTCCAGATGC	604
Db	623	AAGCAGTCCGTGCTTGAACAATGCCAGGATTTCAATTACAGACATCCGCTTCAGATGC	682
QY	605	TGCGGAAGAAGAAGACCACTACTTCGAGCTGTACGAGGCCAGCGTGCGCCGTACAGCG	664
Db	683	CGAGGAAAAAGAAAGACGAGTACTTTGAACCTGTACGAGCATCTGTGGCTAGGTACAGTG	742
QY	665	ACCTGCCCCGAGAAGGCCAATGCA	687
Db	743	ATCTTCTGAAAAAGCAAAATTGA	765

RESULT 12				
AX527890				
LOCUS	AX527890	908 bp	DNA	linear
DEFINITION	Sequence 3	from Patent WO0230965.		PAT 21-NOV-2002
ACCESSION	AX527890			
VERSION	AX527890.1	GI:25172344		
KEYWORDS	.			
SOURCE	Heteractis crispa (leathery sea anemone)			
ORGANISM	Heteractis crispa			
	Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;			
	Stichodactylidae; Heteractis.			
REFERENCE	1			
AUTHORS	Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.			
TITLE	Nucleic acid encoding stichodactylidae chromoproteins			
JOURNAL	Patent: WO 0230965-A 3 18-APR-2002;			
	Clontech Laboratories Inc. (US)			
FEATURES	Location/Qualifiers			
Source	1..908			
	/organism="Heteractis crispa"			
	/mol_type="unassigned DNA"			
	/db_xref="taxon:175771"			

ORIGIN

Query Match	62.9%;	Score 431.8;	DB 6;	Length 908;
Best Local Similarity	77.0%;	Pred. No. 1.4e-46;		
Matches 526;	Conservative 0;	Mismatches 157;	Indels 0;	Gaps 0;

[illegible]



Db	198	TGAGGATTTCATGTACCGAAGGGGCTCCATTACCATTTGCTTCGACATTTTGGCACCGT	257
QY	185	GCTGCGAGTACGGCAGCAGAGACTTTCGTGACCAACACCGCCGAGATCCCGCACTTCTTCA	244
Db	258	GTTGTGAGTACGGCAGCAGAGACTTTCGTGACCAACATACGCGAGAGATTCGCAATTTCTTCA	317
QY	245	AGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAAACCACTACGAGAGACGGCGCA	304
Db	318	AGCAGTCTTTCCTGAAAGGCTTTACTTGGGAAAGAACCAACTATGAAAGATGAGGCA	377
QY	305	TCCTGACCGCCACCAGACACAGCCTTGAGGGCACTGCTGTATCTTCAAGGTGAAG	364
Db	378	TTCTTACTGCTCATCAGGACACAGCCTGAGGGGACTGCTTATATACAAAGTGAAAG	437
QY	365	TGCACGGGACCAACTTCCCGCCGACGGCCCCGTGATGAAAGAACAAAGACGGCGCTGG	424
Db	438	TCCTTGTATCCAAATTTCTGCTGTATGGCCCCGTGATGAAAGAACAAACAGAGATGG	497
QY	425	AGCCGACACCGAGGTGTGTACCCCGAGAACGGCGTGTGCGCGCCGAAACGTGATGG	484
Db	498	AGCCATGCACCTGAGGTGTTTATCCAGATTAATGTGTCTGTGTGAAGTATGTGATGG	557
QY	485	CCCTGAAGTGGCGACCGGCACCTGATCTGCCACCACTACACCACTACCGGAGCAAGA	544
Db	558	CCCTTAAAGTCGTGATCGTCTTGTATCTGCCATCTCTATCTTCTTACAGGTCCAAAG	617
QY	545	AGGCGGTGCGCGCCCTGACCATGCGCGCTTCCACTTCAACGACATCCGGCTCCAGATGC	604
Db	618	AAGCAGTCCGTGCTTGACAATGCCAGATTTCAATTACAGACATCCGCTTCAGATGC	677
QY	605	TGCGGAGAGAGAGACGAGTACTTGTGAGCTGTACGAGCGCCAGCGTGCCCGGTACAGCG	664
Db	678	CGAGGAAAAAGAGAGAGACTTTTGAACGTGACGAAGCATGTGTGCTAGGTACAGTG	737
QY	665	ACCTGCCCCGAGAGGCCCACTGA	687
Db	738	ATCTTCTTGAAAAAGCAATTGA	760
RESULT 13			
AX527912			
LOCUS	AX527912	680 bp	DNA
DEFINITION	Sequence 25 from Patent WO0230965.	linear	PAT 21-NOV-2002
ACCESSION	AX527912		
VERSION	AX527912.1	GI:25172357	
KEYWORDS			
SOURCE	Heteractis crispa (leathery sea anemone)		
ORGANISM	Heteractis crispa		
REFERENCE	1		
AUTHORS	Lukyanov,S.A., Fradkov,A.F., Lukyanov,K.A. and Gurskaya,N.G.		
TITLE	Nucleic acids encoding stichodactylidae chromoproteins		
JOURNAL	Patent: WO 0230965-A 25 18-Apr-2002;		
FEATURES	Clontech Laboratories Inc. (US)		
source	location/Qualifiers		
	1..680		
	/organism="Heteractis crispa"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:175771"		
ORIGIN			
Query Match			
Best Local Similarity 62.5%; Score 429.2; DB 6; Length 680;			
Matches 529; Conservative 0; Mismatches 148; Indels 1; Gaps 1;			
QY	10	GGCCTGCTGAAGAGAGATATGCGCATCAAGATGTACATGAGGGGCAACCGGAC	69
Db	4	GTTTGTGAAAGAAAGTATATGCGCATCAAGATGTACATGAAAGGCAACGTTAATGCCAT	63
QY	70	TACTTCAAGTGCAGAGGCGAGGGCGCAACCCCTTGCCTGCGGCAACCCAGAGCATGAGA	129
Db	64	TATTCAGTGTGAAGAGAGAGGAGACGGCAACCCATTGTGACAGGTACGACAGCATGAGG	123

QY	130	ATCCAGTGACCGAGGGGCCCCCTTGCCCTTGCCCTTGACATCTGCCCCCTGCTGC	189
Db	124	ATTGATGTACCGAAGGGGCTCCATTACCATTTGCTTCGACATTTTGGCAACCGTGTGT	183
QY	190	GAGTACGGCAGCAGACCTTCGTGCCAACACACCGCCGAGATCCCGCACTTCTTCAAGCAG	249
Db	184	GCGTACGGCAGCAGACCTTGTGCCAACATACGGCAGAGATTCCCGATTCTTCAAGCAG	243
QY	250	AGCTTCCCGAGGGCTTCACTTGAGAGAACCACTACGAGAGACGGCGATCTCTG	309
Db	244	TCTTTCCTGAAGGCTTTACTTGGGAAAGAACCAACTATGAAAGATGAGGCAATTCTT	303
QY	310	ACCGCCACCGAGACACAGCCTTGAGGGCACTGCTGTATCTACAAGGTGAAGTGCAC	369
Db	304	ACTGCTCATCAGGACACAAAGCCTGAGGGGAAGTGCCTTATATATACAAGGTGAAGTCC	363
QY	370	GGCACCAACTTCCCGCCGACGGCCCCGTGATGAAAGAACAAAGCGCGCTGGAGCCC	429
Db	364	GGTACCAATTTCTGCTGATGGCCCCGTGATGAAGAA-AAATCAGAGAGATGGAGCCA	422
QY	430	AGCACCGAGTGTGTACCCCGAGAACGGCGTGTGCGCGCCGGAACGTGATGGCCCTG	489
Db	423	AGCACTGAGGTGTTATCCAGAGAAATGTGTCTGTGTGACGTAATGTGATGGCCCTT	482
QY	490	AAAGTGGGCGACCGGCACCTGATCTGCCACCACTACAGCACTACCGGAGCAAGAGCC	549
Db	483	AAAGTCGGTGATCGTCTTGTATCTGCCATCACTATATCTTTACAGAGTCCAAAGAAAGCA	542
QY	550	GTGCGCGCCCTGACCATGCCCCGCTTCCACTTCAACGACATCCGGCTCCAGATGCTGCG	609
Db	543	GTCCGTGCTTGACAATGCCAGATTTCAATTACAGACATCCGCTTCAAGATGCTGAGG	602
QY	610	AAGAGAGAGACGAGTACTTGCAGCTGTACGAGGCGCAGCGTGCCCGGTACAGCACTTG	669
Db	603	AAAAAGAAAGACGAGTACTTTGAACGTGTACGAAGCATCTGTGCTAGGTACAGTATCTT	662
QY	670	CCCCGAAAGGCCCACTGA	687
Db	663	CCTGAAAAAGCAAAATTGA	680

RESULT 14			
AY233273			
LOCUS	AY233273	699 bp	mRNA
DEFINITION	Synthetic construct kindling fluorescent protein mRNA, complete cds.	linear	SYN 02-MAR-2003
ACCESSION	AY233273		
VERSION	AY233273.1	GI:28629492	
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	synthetic construct		
REFERENCE	1		
AUTHORS	Chudakov,D.M., Belousov,V.V., Zaraisky,A.G., Novoselov,V.V., Staroverov,D.B., Zorov,D.B., Lukyanov,S. and Lukyanov,K.A.		
TITLE	Kindling fluorescent proteins for precise in vivo photolabeling		
JOURNAL	Nat. Biotechnol. 21 (2), 191-194 (2003)		
FEATURES	22447893		
source	12524551		
	2 (bases 1 to 699)		
	/organism="synthetic construct"		
	/mol_type="mRNA"		
	/db_xref="taxon:32630"		
	1..699		
	/note="GFP-like protein; capable of photoconversion from nonfluorescent to red fluorescent form; KFP1, KFP-red"		



/codon\_start=1  
/transl\_table=11  
/product="kindling fluorescent protein"  
/protein\_id="AA043187.1"  
/db\_xref="GI:28629493"  
/translation="MASLTETMPFKTTIEGTVNGHCFKCIKGEGNPFEQTQEMKIE  
VIEGGPLPFAFHILSTCMYKSTFKIVSGIPDYFKQSFPFGFTWERTTYEDGGFL  
TAHQDTSLDGLVYKVKILGNNFPADGPMQNKVGRWEPSTEIVYEVDTGLRQSLM  
ALKCPGGRHLTCHLHTYRSKKPASALKMPGFHFEDHRIEIMEEVEKKCYQEAAV  
GRYCDAPSKLGHN"

ORIGIN

Query Match	61.8%;	Score 424.8;	DB 12;	Length 699;
Best Local Similarity	79.3%;	Pred. No. 1.1e-45;		
Matches 530;	Conservative 0;	Mismatches 132;	Indels 6;	Gaps 2;
QY	5	TGAGCGGCTGCTGAAGAGATATGCGCATCAAGATGTACATGAGGGCACCGTGAACG	64	
Db	2	TGGCTTCCTGCTGCTGACCGAGACCATGCTTCAAGACCAACCATGAGGGCAACCGTGAACG	61	
QY	65	GCCACTACTTCAAGTGCAGGGCGAGGGCGACCGCAACCCCTTGCCTGCGCACCCAGAGCA	124	
Db	62	GCCACTGCTTCAAGTGCATCGGCAAGGGCGAGGGCAACCCCTTGAAGGGCACCCAGAGAGA	121	
QY	125	TGAGAAATCCACGTGACCGAGGGCGCCCCCTGCTTGCCTTGCATCTGCGCCCTT	184	
Db	122	TGAAGATGAGGTGATCGAGGGCGCCCCCTGCTTGCCTTGCATCTGCTTCAACT	181	
QY	185	GCTGCGAGTACGGCAGCAGGACCTTGTGACACACACCGCCGAGATCCCGACTTCTTCA	244	
Db	182	CCTGCATGTACGGCTCCAAAGACCTTCATCAAGTACGTGTCCGGCATCCCGACTACTTCA	241	
QY	245	AGCAGAGCTTCCCGGAGGGCTTCACTGGGAGAGAACCAACCACTTACGAGAGCGCGCA	304	
Db	242	AGCAGTCTTCCCGAGGGCTTCACTGGGAGCGCACCACTTACGAGAGCGCGGCT	301	
QY	305	TCCTGACCGCCCAACAGACACCAAGCTGAGGGCAACTGCTTGAATCTACAAAGTGAAG	364	
Db	302	TCCTGACCGCCCAACAGACACCTCTGGAAGCGGCACTGCTGTGTACAAAGTGAAGA	361	
QY	365	TGCACGGCAACCACTTCCCGCGAGCGCCCGTGATGAAGAACAAAGCGGGCGCTGG	424	
Db	362	TCTGTGTAACAACCTTCCCGCGAGCGCCCGTGATGAGAACAAAGTGGCGCTGG	421	
QY	425	AGCCAGCACCGAGGTGTATCCCGAGAACGGCGTGTGCGCGCGGAACGTGATG	484	
Db	422	AGCCCGAACCAGATCGTGTACGAGTGAACGGCGTGTGCGCGGCACTCCCTGATG	481	
QY	485	CCCTGAAG---GTGGCGACCGGCACTGATCTGCCACCACTACACAGCTACCGAGCA	541	
Db	482	CCCTGAAGTGGCCCCGGCGCGGCACTGACCTGCCACCTGACACCACTACCGCTCCA	541	
QY	542	AGAAAGCCCTGCGCGCCCTGACCATGCCCCGCTTCCACTTCCAGAGACCAACCGCATCGA	601	
Db	542	AGAAAGCCCTGCGCGCCCTGAGATGCCCCGCTTCCACTTCCAGAGACCAACCGCATCGA	601	
QY	602	TGCTGCGGAA---GAAGAGACGAGTACTTGCAGCTGTACGAGGCCAGCGTGCGCCGT	658	
Db	602	TGCTGCGGAGGTGAGAGAGGCAAGTGTCTACAAAGCAGTACGAGCGCGCTGCGCGCT	661	
QY	659	ACAGCGAC 666		
Db	662	ACTGCGAC 669		

RESULT 15  
AX824739  
LOCUS AX824739 707 bp DNA linear PAT 11-DEC-2003  
DEFINITION Sequence 21 from Patent WO02068459.  
ACCESSION AX824739  
VERSION AX824739.1 GI:39750599  
KEYWORDS  
SOURCE synthetic construct

ORGANISM synthetic construct  
REFERENCE 1  
AUTHORS  
TITLE Non aggregating fluorescent proteins and methods for using the same  
JOURNAL Patent: WO 02068459-A 21 06-SEP-2002;  
FEATURES  
source location/Qualifiers  
1.707  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="non-aggregating mutant"

ORIGIN

Query Match	60.8%;	Score 417.8;	DB 6;	Length 707;
Best Local Similarity	79.1%;	Pred. No. 9e-45;		
Matches 523;	Conservative 0;	Mismatches 132;	Indels 6;	Gaps 2;
QY	12	CCTGCTGAAGAGAGATGCGCATCAAGATGTACATGAGGGCACCGTGAACGGCACTA	71	
Db	12	CCTGCTGACCGAGACCATGCTTCAAGACCAACCATGAGGGCAACCGTGAACGGCACTA	71	
QY	72	CTTCAAGTGCAGGGCGAGGGCGACCGCAACCCCTTGCCTGCGCACCCAGAGCATGAGAT	131	
Db	72	CTTCAAGTGCACCGGCAAGGGCGAGGGCAACCCCTTGAAGGGCACCCAGAGATGAGAT	131	
QY	132	CAACGTGACCGAGGGCGCCCCCTGCTTGCCTTGCATCTGCGCCCTGCTGCA	191	
Db	132	CGAGGTGATCGAGGGCGCCCCCTGCTTGCCTTGCATCTGCTTGCATCTGCTGCA	191	
QY	192	GTAAGGAGCAGACCTTGTGACACCAACCGCCGAGATCCCGACTTCTTCAAGCAGAG	251	
Db	192	GTAAGGCTCCAAAGCCTTCATCAAGTACGTGTCCGGCATCCCGACTACTTCAAGCAGTC	251	
QY	252	CTTCCCGGAGGGCTTCACTGGGAGAGAACCAACCACTTACGAGAGCGCGCATCTGAC	311	
Db	252	CCTCCCGAGGGCTTCACTGGGAGCGCACCACTTACGAGAGCGCGCTTCTGAC	311	
QY	312	CGCCCAACAGGACACCACTGAGGGCAACTGCTGTATCTACAAAGTGAAGTGAACGG	371	
Db	312	CGCCCAACAGGACACCTCTGAGCGGCACTGCTGTGTATCAAGTGAAGTCTTGG	371	
QY	372	CACCAACTTCCCGCGAGCGCCCGTGATGAAGAACAAAGCGGGCGTGGAGCCAG	431	
Db	372	CAACAACCTTCCCGCGAGCGCCCGTGATGAGAACAAAGCGGGCGTGGAGCCCTC	431	
QY	432	CACGAGGTGTATCCCGAGAACGGCGTGTGCGCGCGGAACGTGATGCGCTGAA	491	
Db	432	CACCGAGATGTGTACAGGTGAGACGGCGTGTGCGCGGCACTCCCTGATGCGCTGGA	491	
QY	492	G---GTGGCGACCGGCACTGATCTGCCACCACTACACAGCTACCGGAGCAAGAGCC	548	
Db	492	GTCGCCCGGCGGTCGCACTGACCTGCCACCTGACACCACTACCGCTCAAGAGCC	551	
QY	549	CGTGCGCGCCCTGACCATGCCCCGCTTCCACTTACCGACATCCGGCTCCAGATGCTCG	608	
Db	552	CGCTCCGCGCTGAAGATGCCCCGCTTCCACTTGAAGAACCAACCGCATGAGATCTGGA	611	
QY	609	GAA---GAAGAGACGAGTACTTGCAGCTGTACGAGGCCAGCGTGCCCGGTACAGCGA	665	
Db	612	GGAGGTGAGAGAGGCAAGTGTCTACAAAGCAGTACGAGGCGCGGTGCGCTGCGGA	671	
QY	666	C 666		
Db	672	C 672		

Search completed: July 15, 2004, 13:03:04  
Job time : 3028 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2004, 11:17:23 ; Search time 82 Seconds  
(without alignments)  
4649.406 Million cell updates/sec

Title: US-09-976-673-11

Perfect score: 687

Sequence: 1 atcggtgagcgctgtgtgaa.....tgcccgagaagccaactga 687

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 segs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	253.2	36.9	720	4	US-09-839-650-1 Sequence 1, Appli
2	176.6	25.7	678	4	US-09-459-956-6 Sequence 6, Appli
3	153.8	22.4	801	4	US-09-459-956-7 Sequence 7, Appli
4	145.4	21.2	699	4	US-09-459-956-5 Sequence 5, Appli
5	121.6	17.7	1104	3	US-09-277-716-30 Sequence 30, Appli
6	121.6	17.7	1104	4	US-09-609-161B-30 Sequence 30, Appli
7	121.6	17.7	1279	3	US-09-277-716-31 Sequence 31, Appli
8	121.6	17.7	1279	4	US-09-172-063-14 Sequence 14, Appli
9	110.6	16.1	720	3	US-09-316-919-15 Sequence 15, Appli
10	110.6	16.1	720	4	US-09-316-919-15 Sequence 15, Appli
11	110.6	16.1	720	4	US-09-602-641-14 Sequence 14, Appli
12	109.8	16.0	690	4	US-09-459-956-2 Sequence 2, Appli
13	109.8	15.9	717	4	US-09-513-783A-43 Sequence 43, Appli
14	109.8	15.9	720	3	US-08-911-825-3 Sequence 3, Appli
15	109.8	15.9	720	3	US-08-974-737-3 Sequence 3, Appli
16	109.8	15.9	720	3	US-08-706-408-3 Sequence 3, Appli
17	109.8	15.9	720	3	US-09-094-359-9 Sequence 9, Appli
18	109.8	15.9	720	3	US-09-172-063-12 Sequence 12, Appli
19	109.8	15.9	720	3	US-09-172-063-17 Sequence 17, Appli
20	109.8	15.9	720	4	US-09-465-142-3 Sequence 3, Appli
21	109.8	15.9	720	4	US-09-316-919-13 Sequence 13, Appli
22	109.8	15.9	720	4	US-09-316-919-18 Sequence 18, Appli
23	109.8	15.9	720	4	US-09-575-847-3 Sequence 3, Appli
24	109.8	15.9	720	4	US-09-602-641-12 Sequence 12, Appli
25	109.8	15.9	720	4	US-09-602-641-17 Sequence 17, Appli
26	109.8	15.9	720	4	US-09-704-463-9 Sequence 9, Appli
27	109.8	15.9	768	3	US-09-172-063-28 Sequence 28, Appli

28	109	15.9	768	3	US-09-172-063-33	Sequence 33, Appl
29	109	15.9	768	4	US-09-602-641-28	Sequence 28, Appl
30	109	15.9	768	4	US-09-602-641-33	Sequence 33, Appl
31	109	15.9	972	3	US-09-172-063-30	Sequence 30, Appl
32	109	15.9	972	4	US-09-602-641-30	Sequence 30, Appl
33	109	15.9	1623	4	US-09-513-783A-33	Sequence 33, Appl
34	109	15.9	2439	4	US-09-513-783A-3	Sequence 3, Appli
35	109	15.9	2439	4	US-09-513-783A-5	Sequence 5, Appli
36	109	15.9	3171	4	US-09-513-783A-31	Sequence 31, Appl
37	109	15.9	4833	4	US-09-513-783A-21	Sequence 21, Appl
38	107.8	15.7	720	3	US-09-316-919-19	Sequence 19, Appl
39	107.8	15.7	720	4	US-09-316-919-18	Sequence 18, Appl
40	107.8	15.7	768	3	US-09-172-063-34	Sequence 34, Appl
41	107.8	15.7	768	4	US-09-602-641-34	Sequence 34, Appl
42	107.8	15.7	720	3	US-09-094-359-5	Sequence 5, Appli
43	107.4	15.6	720	4	US-09-704-463-5	Sequence 5, Appli
44	107.4	15.6	762	1	US-08-532-390-40	Sequence 40, Appl
45	107.4	15.6				

ALIGNMENTS

RESULT 1  
US-09-839-650-1  
; Sequence 1, Application US/09839650  
; Patent No. 6645761  
; GENERAL INFORMATION:  
; APPLICANT: Stratagene  
; TITLE OF INVENTION: Humanized Polynucleotide Sequence Encoding Renilla Mulleri Green  
; Patent No. 6645761  
; TITLE OF INVENTION: Fluorescent Protein  
; FILE REFERENCE: 25436/1755  
; CURRENT APPLICATION NUMBER: US/09/839, 650  
; CURRENT FILING DATE: 2001-04-19  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 1  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Humanized R. mulleri polynucleotide  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(720)  
; OTHER INFORMATION: Humanized DNA sequence  
; US-09-839-650-1

Query Match 36.9%; Score 253.2; DB 4; Length 720;  
Best Local Similarity 62.0%; Pred. No. 1.9e-43;  
Matches 418; Conservative 0; Mismatches 253; Indels 3; Gaps 1;

QY	2	TGCTGAGCGGCGCTGCGAAGAGATGCGCATCAAGATGTACATGAGGGCAACCTGA	61
DB	20	TGAAGAACACCTGCTGCAGAGAGTGATGAGCTTCAAGGTGAACCTGAGGGCATCGTGA	79
QY	62	ACGGCCACTACTTCAAGTGCAGGGCGAGGGCGAAGCAACCCCTTCCGGGCAACCCAGA	121
DB	80	ACAACACGTTGTTCAACATGAGAGGGCTGCGCAAGGCAACATCTGTTGGCAACAGC	139
QY	122	GCATGAGATCCACGTGACCGAGGGCGCCCTTGCCTTGCCTTGCACATCCTGGCCC	181
DB	140	TGCTGACAGATCCGCGTGAACCAAGGGCGCCCTTGCCTTGCCTTGCACATCCTGAGCC	199
QY	182	CCTGCTGCGAGTACGGGACGACGACCTTCTGTGACCAACCGCCGAGATCCCGACTTCT	241
DB	200	CCGCTTCCAGTACGGGACCGACCTTCAACCAAGTACCCCAACGACATCAGCACTACT	259
QY	242	TCAAGCAGAGCTTCCCGAGGGGCTTCACTTGGGAGAGAACAACCAACCTTACGAGCGCG	301
DB	260	TCAATCAGAGCTTCCCGCGGCTTCAATGATACGAGCGCAACCTTGGCTTACGAGCGCG	319
QY	302	GCATCTGACCGGCCACGAGACCAACGACCTTGAAGGGCAACTGCTGATCTACAAGGTGA	361

```
Db      320 GCCTGCTGAGATCCGACGCGCATCAACCTGATCGAGGACAAAGTTGCTAACCGCGTGG 379
QY      362 AGGTGCACGCGCACTTCCCGCGCGGACCGCGCGGTGATGAAGAACAGAGCGGCGGT 421
Db      380 AGTACAAAGGCGACACTTCCCGACGACCGCGCGGTGATGCAAGAACCATCTCGGCA 439
QY      422 GGGAGCCCAAGACCGAGTGGTGTACCCCGAGAACGCGCGTGTGCGCGCGGAACGTGA 481
Db      440 TCGAGCCCAAGCTTCGAGCGCATGTACATGAACAACGCGCGTGTGCGCGAGGTGATCC 499
QY      482 TGGCCCTGAAGGTGGGCGGACCGGACCTGATCTGCCACCACTACACAGCTACCGGAGCA 541
Db      500 TGGTGTACAAAGCTGAACAGCGGCGCAAGTACTACAGCTGCCACATGAAGACCCCTGATGAAGA 559
QY      542 AGAAGCGCGTGGCGCGCGCTGACCATGCGCGGCTTCCACTTCACCGGACATCCGCGTCCAGA 601
Db      560 GCAAGGCGGTG--GTGAAGAGAGTCCCTCTCTACCACTTCATCCAGACCGCGCTGGAGA 616
QY      602 TGCTGCGGAAGAGAAGAGAGAGAGTACTTTCAGCTGTACGAGGCCAGCGGTGCGCGGTACA 661
Db      617 AGACCTAAGTGAAGAGAGAGCGCGCTGTGTGAGAGACAGAGACCGCCCATCGCCAGATGA 676
QY      662 GCGACCTGCGCGAG 675
Db      677 CCAGCATCGGCAAG 690
```

## RESULT 2

```
US-09-459-956-6
; Sequence 6, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Discosoma sp
US-09-459-956-6
```

```
Query Match      25.7%; Score 176.6; DB 4; Length 678;
Best Local Similarity 60.2%; Pred. No. 8.7e-28;
Matches 293; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
```

```
QY      17 TGAGGAGAGATGCGCATCAAGATGTATGAGGGGACCGGTGAACGCGCACTACTTCA 76
Db      23 TCAAGAGATTGATGAGGTTTAAAGTTTCGATGGAAGAACGGTCAATGGGACGAGTTTG 82
QY      77 AGTGCAGAGGCGAGGGCGACGCAACCCCTTCCCGGCGACCGAGAGCATGAATCCACG 136
Db      83 AATAAGAGGCGAAGAGAGGGGCGCCATACGAAGGCCACAATACCGTAAAGCTTAAGG 142
QY      137 TGACCGAGGGGCGCGCGCGCTTCCGCTTGCATCCGCGCGCGCGCGCGCGAGTACG 196
Db      143 TAAACCAAGGGGGGACCTTTGCAATTTGCTTGGGATATTTGTCAACCAATTTCAAGTATG 202
QY      197 GCAGCAGAGACCTTCTGTGACACACACCGCGAGATCCCGGACTTCTTCAAGCAGAGCTTC 256
Db      203 GAAGCAGAGTATATGTCAAGCACCTGCGGACATACCAAGTATATAAAAGCTGTCAATTC 262
```

```
QY      257 CCGAGGCGCTTCACTTGGAGAGAACCACTACGAGAGCGCGGCGCATCTGACCGGCC 316
Db      263 CTGAAGGATTTAATGGGAAGGGGTCAATGACTTTGAAGACGGTGGCGTCTTACTGTAA 322
QY      317 ACCAGACACCAAGCGCTGAGGCGCACTGCTGATCTTCAAGGTGAAGGTGACCGCACCA 376
Db      323 CCCAGGATTCAGTTTGACAGATGCGTGTTCATCTACAGGTCAAGTTCATTGGCGTGA 382
QY      377 ACTTCCCGCGGACGCGCGCGGTGATGAAGAACAGAGCGCGCGGTGAGCGGACGACCG 436
Db      383 ACTTCTTCCGATGACCTGTTATGCAAAAGAGACATGGGCTGGGAAGCCAGCACTG 442
QY      437 AGGTGTGTACCCCGAAGACGGCGTGTGTGCGCGCGCGGAACTGATGCTCTGAAGGTGG 496
Db      443 AGCGTTGTATCTCGTGATGCGGTGTGAAGAGAGATTCATAAGGCTCTGAAGCTGA 502
QY      497 GCGACCG 503
Db      503 AAGACCG 509
```

## RESULT 3

```
US-09-459-956-7
; Sequence 7, Application US/09459956
; Patent No. 6342379
; GENERAL INFORMATION:
; APPLICANT: Tsien, Roger Y.
; APPLICANT: Gonzalez, III, Jesus E.
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY
; FILE REFERENCE: REGEN1290-4
; CURRENT APPLICATION NUMBER: US/09/459,956
; PRIOR FILING DATE: 1999-12-13
; PRIOR APPLICATION NUMBER: 08/765,860
; PRIOR FILING DATE: 1999-05-08
; PRIOR APPLICATION NUMBER: 08/481,977
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: PCT/US96/09652
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 801
; TYPE: DNA
; ORGANISM: Clavularia sp
US-09-459-956-7
```

```
Query Match      22.4%; Score 153.8; DB 4; Length 801;
Best Local Similarity 58.4%; Pred. No. 4.1e-23;
Matches 269; Conservative 0; Mismatches 192; Indels 0; Gaps 0;
```

```
QY      10 GGCCTGTAAGAGAGATATGCGCATCAAGATGTATGAGGGGACCGGTGAACGCGCAC 69
Db      130 GGTGTGATTAACCAACATGAAGATTTAAGCTGAAGATGAAGAAATGTAAACGCGCAT 189
QY      70 TACTCAAGTGCAGGGGCGAGGGCGGCAACCCCTTGCCTTGCATTCCTGCGCGCGCGCAT 129
Db      190 GCTTTGTGATCGAAGAGAGAGAGAAAGGCTTACGATGGGACACACACTTTAAAC 249
QY      130 ATCCAGTGAACCGAGGGCGCGCGCGCTTGCCTTGCATTCCTGCGCGCGCGCATCTGC 189
Db      250 CTGGAAGTGAAGAGAGGTGCGCTTGCCTTTTCTTACGATATCTGTCAAACGCGTTC 309
QY      190 GAGTACGGCAGGAGACCTTGTGCAACCAACCGCGGAGATCCCGACTTCTTCAAGCAG 249
Db      310 CAGTACGGAACAGACGATTTGCAAAATACCCAGAGATATAGCAGACTATTTCAAGCAG 369
QY      250 AGCTTCCCGGAGGCTTCACTGGGAGAGAACCAACCACTTACGAGAGCGCGCATCTGC 309
Db      370 TCGTTTCCGAGGATATTTCTGGGAAAGAACCATGACTTTTGAAGACAAAGGCTTGTG 429
QY      310 ACCGCCACCAAGACACAGCCTGAGGGGCACTGCTGATCTTCAAGGTGAAGGTGAC 369
```



Db 430 AAATGAAAAAGTGACATAGCATGAGGAAGACTCCCTTATCTATGAAATTGTTTGAAT 489  
QY 370 GGACCAACTTCCCGCGGACGGCCCGTGATGAAGAACAGCGCGCTGGAGCCC 429  
Db 490 GGATGAACTTCTCCCAATGGTCCGGTTATGCAGAAAAAACTTTGAAGTGGAAACA 549  
QY 430 AGACCGAGGTGTGTACCCCGAAGACGGCGTGTGCGG 470  
Db 550 TCCACTGAGATTATGTACGTGCGTGATGAGTGTGTCGG 590

RESULT 4

US-09-459-956-5  
; Sequence 5, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; TITLE OF INVENTION: OPTICAL METHODS  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; CURRENT FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; PRIOR FILING DATE: 1996-06-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 699  
; TYPE: DNA  
; ORGANISM: *Discozoma striata*  
US-09-459-956-5

Query Match 21.2%; Score 145.4; DB 4; Length 699;  
Best Local Similarity 56.3%; Pred. No. 2.1e-21;  
Matches 272; Conservative 0; Mismatches 211; Indels 0; Gaps 0;

QY 14 TCGTGAAGAGATATGCGCATCAAGATGTACATGAGGGCAACCGTGAACGCCACTACT 73  
Db 20 TGATCAAGGAAGAATGTGATCGATCTTCATCTGAAAGAACGTTCAATGGCACACTACT 79  
QY 74 TCAAGTCGAGGGCGGAGGCGGCAACCCCTTCGCCGCAACCCAGACATGAGAAATCC 133  
Db 80 TTGAATAAAGGCAAGGAAGAAAGACAGCCTAATGAAGGCAACCAATACGTCACGCTCG 139  
QY 134 ACGTGACCGAGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 193  
Db 140 AGGTTACCAAGGGGTGACCTCTGCACTTTGGTTGGCATATTTTGGCCCAATTTTCAGT 199  
QY 194 ACGGACGAGGACCTTCGTGCACCAACCGCGAGATCCCGACTTCTCAAGCAGAGCT 253  
Db 200 ATGGAACAAGGCATTTGTCACCACTGACAAACATACATGATTAATCTAAAGCTGTGAT 259  
QY 254 TCCCGAGGGCTTCACCTGGAGAGAAACCAACCTACGAGGACGGCGCATCTGACCG 313  
Db 260 TTCCGGAGGATATACATGGGAACGCTCCATGCACCTTTGAAGACGGTGGCTGTGTGTA 319  
QY 314 CCGACGAGACACAGCCTGAGGGCAACTGCCTGATCTAAGAGTGAAGTGCACGCA 373  
Db 320 TCACCAATGATATCAGTTTGACAGGCACTGTTCTACTAGACATCAAGTTCACTGGCT 379  
QY 374 CCAACTTCCCGCGGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 433  
Db 380 TGAACCTTCTCCAAATGACCCGTTGTGCAAGAAAGACAACTGGCTGGAAACCGAGCA 439  
QY 434 CCGAGGTGTGTACCCCGAGAACGGCGGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 493  
Db 440 CTGAGCGTTTGTATCTCGTGATGTGTGTGATAGAGACATCCATCATGCTCTGACAG 499

QY 494 TGG 496  
Db 500 TTG 502

RESULT 5

US-09-277-716-30  
; Sequence 30, Application US/09277716A  
; Patent No. 6232107  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.  
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE  
; CURRENT APPLICATION NUMBER: US/09/277,716A  
; CURRENT FILING DATE: 1999-03-26  
; EARLIER APPLICATION NUMBER: 60/102,939  
; EARLIER FILING DATE: 1998-10-01  
; EARLIER APPLICATION NUMBER: 60/089,367  
; EARLIER FILING DATE: 1998-06-15  
; EARLIER APPLICATION NUMBER: 60/079,624  
; EARLIER FILING DATE: 1998-03-27  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 30  
; LENGTH: 1104  
; TYPE: DNA  
; ORGANISM: *Ptilosarcus gurneyi*  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (34)..(747)  
; FEATURE:  
; OTHER INFORMATION: *Ptilosarcus Green Fluorescent Protein (GFP) (insert A)*  
US-09-277-716-30

Query Match 17.7%; Score 121.6; DB 3; Length 1104;  
Best Local Similarity 54.2%; Pred. No. 1.7e-16;  
Matches 247; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 16 CTGAAGAGAGATATGCGCATCAAGATGTACATGAGGGCAACCGTGAACGCCACTACTTC 75  
Db 64 CTGAAGAGATTATGTCGCAAAAGCTAGCGTTGAAGAAATCGTAACAATCAGCTTTT 123  
QY 76 AAGTGCAGGGCGGAGGCGGCAACCCCTTCGCCGCAACCCAGACATGAGAAATCCAC 135  
Db 124 TCCATGGAAGATTGGAAGAAAGCAATGTATATTGGAACCAATTGATGCAATCCGG 183  
QY 136 GTGACCGAGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 195  
Db 184 GTTACAAAGGAGGTCCGTTGCCATTTCGCTTTCGATATGTTTCCATAGCTTCCAAATAC 243  
QY 196 GGACGACGAGACCTTCGTGCACCAACCGCGGAGATCCCCGACTTCTTCAAGCAGAGCTTC 255  
Db 244 GGAATCGCACTTTCACGAATAACCGAGACGACATTCGGACTACTTGTTCATATTCATTC 303  
QY 256 CCGGAGGCTTCACCTGGAGAGAACCAACCACTACGAGGACGGCGCATCTGACCGCC 315  
Db 304 CCGGCTGATTTTTCAGAAAGAAATCTACGCTTTGAAGATGGCGCATGTTGACATT 363  
QY 316 CACCAAGACACAGCCTGAGGGCAACTGCCTGATCTACAAGGTGAAGTGCACGCAAC 375  
Db 364 CGTTCAATATTAAGTTTGAAGATGATTAAGTTCCACTACAAAGTGAGTATAGAGCAAC 423  
QY 376 AACTTCCCGCGGACGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 435  
Db 424 GGTTCCTAGTAACGACCGCGTGATGCAAAAAGCCATCCTCGCATGAGCAATCGTTT 483  
QY 436 GAGGTGTGTACCCCGAGAACGGCGGTGCTGTGCGGC 471  
Db 484 GAGGTGTGTACATGAACAGCGGCGGCTTCTGTGCGGC 519

Fri Jul 16 08:44:54 2004

us-09-976-673-11.rn1

Page 4

```

RESULT 6
US-09-609-161B-30
; Sequence 30, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
; APPLICANT: Bryan, Bruce
; APPLICANT: Szent-Gyorgyi, Christopher
; APPLICANT: PROLUME, LTD.
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 1104
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (34)..(747)
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (insert A)
US-09-609-161B-30

```

Query Match	17.7%;	Score 121.6;	DB 4;	Length 1104;
Best Local Similarity	54.2%;	Pred. No. 1.7e-16;		
Matches 247;	Conservative	0;	Mismatches 209;	Indels 0;
				Gaps 0;

QY 16 CTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCCTGAACGGCCACTACTTC 75  
 Db 64 CTGAAAGAGATTATGTGCGCAAAAGCTAGCGTTGAAGGAATCGTGAACAATCAGCTTTT 122  
 QY 76 AAGTCGAGGGCGAGGGCGCAGCGCAACCCCTTCGCCGGCACCAGCATGAAGATCCAC 135  
 Db 124 TCCATGGAAGGATTGTGAAAAGCCAAATGATTATTGTGAAAACCAATTGATGCAAAATCCGG 183  
 QY 136 GTGACCGAGGGCGCCCCCTGCGCTTCGCTTCGACATCCTGGCCCCCTGCGAGTAC 199  
 Db 184 GTTACAAAGGAGAGTCCCGTTCCTTCGCTTCGATATTGTTTCATAGCTTCCAAATAC 243  
 QY 196 GGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTC 255  
 Db 244 GGAATCGCACTTTCACGAAATACCCAGACGACATTGCGGACTACTTGTTCATCATTC 303  
 QY 256 CCCGAGGCTTCACCTGGGAGAGAACCAACCACTACGAGGACGGCGCATCTGACCGCC 315  
 Db 304 CCGGCTGATTTTCTACGAAAGAATCTACGCTTGAAGATGGCGCCATTGTTGACATT 363  
 QY 316 CACCAAGACACCAAGCTGAGGGCAACTGCTGATCTACAAAGTGAAGTGCAAGGCACC 375  
 Db 364 CGTTAGATATTAAGTTTGAAGATGATTAAGTTCCACTCAAAAGTGAAGTATGAGGCAAC 423  
 QY 376 AACTTCCCCCGCAGCGCCCCGTGATGAAGAACAAGAGCGGGCGCTGGAGCCCAAGCACC 435  
 Db 424 GGTTCCTTAGTAACGGAACCGGTATGCAAAAAGCCATCTCGCATGAAGCCATCGTTT 483  
 QY 436 GAGGTGCTGTAACCCGAGAAACGGCGTGTGCGC 471  
 Db 484 GAGGTGCTTACATGAACAGCGGCGTTCTGTGGCG 519

RESULT 7  
US-09-277-716-31

```

: Sequence 31, Application US/09277716A
: Patent No. 6232107
: GENERAL INFORMATION:
: APPLICANT: Bryan, Bruce
: APPLICANT: Szent-Gyorgyi, Christopher
: APPLICANT: PROLOME, LTD.
: TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
: CURRENT APPLICATION NUMBER: US/09/277,716A
: CURRENT FILING DATE: 1999-03-26
: EARLIER APPLICATION NUMBER: 60/102,939
: EARLIER FILING DATE: 1998-10-01
: EARLIER APPLICATION NUMBER: 60/089,367
: EARLIER FILING DATE: 1998-06-15
: EARLIER APPLICATION NUMBER: 60/079,624
: EARLIER FILING DATE: 1998-03-27
: NUMBER OF SEQ ID NOS: 32
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 31
: LENGTH: 1279
: TYPE: DNA
: ORGANISM: Ptilosarcus guineyi
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (7)..(720)
: FEATURE:
: OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (Insert B)
US-09-277-716-31

```

Query Match	17.7%;	Score 121.6;	DB 3;	Length 1279;
Best Local Similarity	54.2%;	Pred. No. 1.7e-16;		
Matches 247;	Conservative	0;	Mismatches 209;	Indels 0;
				Gaps 0;

QY	16	CTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCAACCGGCACTACTTC	75
Db	37	CTGAAAGAGATTATGTGCGCAAAAGCTAGCGTTGAAGGAATCGTGAACATCAACGTTTTT	96
QY	76	AAGTGGAGGGCGAGGGCGAAGCAACCCCTTTCGCGCGCAACCCAGAGCATGAAATCCAC	13
Db	97	TCCATGGAAGGATTGTGAAAAGGCAATGTAATTAATTGGAACCAATTGATGCAAAATCCGG	15
QY	136	GTGACCGGAGGGCGCCCCCTGCCCCCTTTCGCTTCGACATCTGGCCCCCTGCTGGAGTAC	19
Db	157	GTTACAAAGGAGGATCCGTTGCCATTGCTTTTCGACATGTGTTCCATAGCTTTCCAATAC	21
QY	196	GGCAGCAGGACCTTCGTGCAACACACCGCGAGATCCCGCACTTCTTCAAGCAGAGCTTC	25
Db	217	GGGAATCGCACTTTCACGAATATACCGACAGACATTGGCGACTACTTGTTCAAATCATTT	27
QY	256	CCCGAGGGCTTCACTGCGGAGAGAACCAACCTACGAGAGCGCGGCATCTGACCGCC	31
Db	277	CCGGCTGGATTTTTTCTACGAAGAATACTACGCTTTGAAGATGGCGCAATTGTGACATT	33
QY	316	CACCAAGACACACCGCTGGAGGGCACTGCTGTATCTACAAGGTGAAGGTGCACGGCACCC	37
Db	337	CGTTCAAGATATAAGTTTGAAGATGATTAAGTTCCACTACACAAAGTGAAGTATAGAGGCAAC	39
QY	376	AACTTCCCCGCGCGACCGCCCGTGTATGAAGACAAAGCGCGGCTGGAGCCCAAGCACCC	43
Db	397	GGTTTCCCTAGTAAAGGACCGCCGTGATGCAAAAAGCCATCCTCGGCATGAGCCATCGTTT	45
QY	436	GAGGTGTTATCCCGAGAACGGCGTGTGTGCGGC	471
Db	457	GAGGTGTTACATGAACAGCGCGCTTCTGTGTGGGC	492

RESULT 8  
US-09-609-161B-31  
; Sequence 31, Application US/09609161B  
; Patent No. 6436682  
; GENERAL INFORMATION:  
; APPLICANT: Bryan, Bruce  
; APPLICANT: Szent-Gyorgyi, Christopher  
; APPLICANT: PROLUME, LTD.

```
; TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE LUC
; TITLE OF INVENTION: AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS, HIG
; TITLE OF INVENTION: SCREENING AND NOVELTY ITEMS
; FILE REFERENCE: 24729-121B
; CURRENT APPLICATION NUMBER: US/09/609,161B
; CURRENT FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 09/277,716
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/102,939
; PRIOR FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/089,367
; PRIOR FILING DATE: 1998-06-15
; PRIOR APPLICATION NUMBER: 60/079,624
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 1279
; TYPE: DNA
; ORGANISM: Ptilosarcus gurneyi
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(720)
; OTHER INFORMATION: Ptilosarcus Green Fluorescent Protein (GFP) (Insert B)
US-09-609-161B-31
```

Query Match 17.7%; Score 121.6; DB 4; Length 1279;  
Best Local Similarity 54.2%; Pred. No. 1.7e-16;  
Matches 247; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

```
QY 16 CTGAAGAGAGTATGCGCATCAAGATGTACATGAGGGCACCCTGAACGGCCACTACTTC 75
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 37 CTGAAGAAGATATATGTCGGCAAAAGCTAGCGTTGAAGAAATCGTAACATCACGTTT 96
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 76 AAGTCGAGGGCGGAGGGCGCAACCCCTTCGCCGGCACCCAGAGCATGAGAATCCAC 135
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 TCCATGGAAGGATTGGAAAAGGCAATGATTTATTTGAAACCAATTGATGCAAAATCCGG 156
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 136 GTGACCGAGGGCGCCCCCTGCCCTTCGCTTCGACATCTGGCCCCCTGTCGAGTAC 195
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 GTTACAAAGGAGGTCCGTTGCCATTCGCTTCGACATGTTTCCATAGCTTTCCAAATAC 216
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 196 GGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCGACTTCTTCAAGCAGACTTC 255
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 GGGAAATCGCACTTTCACGAATACCCAGACGACATTCGGGACTACTTGTCAATCAATT 276
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 256 CCGAGGGCTTCACTGGAGAGAACCAACCACTACGAGACGGCGGCATCCTGACGCC 315
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 277 CCGGCTGATTTTCTACGAAGAATCTACGCTTTGAAGATGGCCCATTTGTGACATT 336
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 316 CACCAAGACACCAACCGCTGAGGGCACTGCCTGATCTACAAGTGAAGTGCACGGCAAC 375
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 CGTTCAGATATAGTTTAGAAGATGATAGTTCCACTACAAAGTGAGTATAGAGGCAAC 396
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 AACTTCCCCCGCGACGGCCCCGTGATGAAGAACAAAGAGCGGCGCTGGAGCCCGACACC 435
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 397 GGTTCCTAGTAAAGGACCGCGTATGCAAAAAGCCATCCTCGGCATGAGCCATCGTTT 456
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 436 GAGGTGCTGTACCCCGAGAACCGCGTGTGTGCGGC 471
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 457 GAGGTGCTCTACATGAACACGCGGCTTCTGTGTGGC 492
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 9  
US-09-172-063-14  
; Sequence 14, Application US/09172063  
; Patent No. 6150176  
; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Miyawaki, Atsushi  
; APPLICANT: Llopis, Juan  
; APPLICANT: Wachter, Rebekka M.  
; APPLICANT: Remington, S. James

```
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR
; TITLE OF INVENTION: MEASURING THE PH OF A BIOLOGICAL SAMPLE
; FILE REFERENCE: 07257/071001
; CURRENT APPLICATION NUMBER: US/09/172,063
; CURRENT FILING DATE: 1998-10-13
; EARLIER APPLICATION NUMBER: 09/094,359
; EARLIER FILING DATE: 1998-06-09
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 720
; TYPE: DNA
; ORGANISM: Aequorea victoria
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: EYFP-V68L/Q69K
US-09-172-063-14
```

Query Match 16.1%; Score 110.6; DB 3; Length 720;  
Best Local Similarity 55.2%; Pred. No. 2.8e-14;  
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;

```
QY 13 CTGCTGAAGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTGAACGGCCACTAC 72
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 22 CTGTTACCGGGGTGTGCTCCATCTGTGAGCTGACGGCGACGTAACGGCCACAAG 81
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 73 TTCAAGTCCGAGGGCGGAGGGCGGACCGCAACCCCTTCGCCGGCACCCAGAGCATGAATC 132
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 82 TTCAGCGTGTCCGGCGGAGGGCGGAGGGCGATGCCACTTACGGCAAGCTGACCTGAAGTTC 141
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 133 CACGTACCGAGGGCGCCCCCTGCCCTTCGCTTGACATCTGGCCCCCTGCTGCGAG 192
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 142 ATCTGCACCAC--CGGCAAGCTGCCGCTGCGCCCAACCCTGTGACCACTTCGGC 198
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 193 TACGGCAGCAGACCTTCGTGCACCAACCAACCGCCGAGAT-----CCCCGACTTCTCAAG 246
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 TACGGCTGAAGTGCTTCGCCCGCTACCCCGACCATGAAGCAGCAGCATCTTCAAG 258
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 247 CAGAGCTTCCCGAGGGCTTCACTGAGAGAACCAACCTACGAGAGCGGCGCATC 306
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 TCCGCCATGCCCGAAGGCTACGTCAGAGCGGCAACCATCTTCTCAAGCAGCAGCAAC 318
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 307 CTGACCGCCCAACCAAGACCAACGCTGAGGGGCACTGCCTGATCTACAAGTGAAGTG 366
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 TACAAGACCCGCGCCGAGGTGAAGTTGAGGGCGCACACCTGTGAACCGCATGAGCTG 378
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 367 CACGGCACCACCTTCCCGCCGACGCGCCCGTGTGAAGAACAAAGCGGCGCTGGAG 426
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 AAGGCATGACTTCAAGAGAGACGGCAACATCTGGGGCACAAAGCTGAGTACAATAC 438
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 427 CCCAGCAGGAGTGTGTACCCCGAGAACGGCGTGTGCGGCGGAACTGA 481
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 439 AACAGCACCAACGTATATATCATGCGCGACAGCAAGAAAGCGCATCAAGGTGA 493
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 10  
US-09-316-919-15  
; Sequence 15, Application US/09316919  
; Patent No. 6469154  
; GENERAL INFORMATION:

; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Baird, Geoffrey  
; TITLE OF INVENTION: FLUORESCENT PROTEIN INDICATORS  
; FILE REFERENCE: 07257/073001  
; CURRENT APPLICATION NUMBER: US/09/316,919  
; CURRENT FILING DATE: 1999-05-21  
; NUMBER OF SEQ ID NOS: 63  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Aequorea victoria

US-09-316-919-15

Query Match 16.1%; Score 110.6; DB 4; Length 720;  
Best Local Similarity 55.2%; Pred. No. 2.8e-14;  
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;

QY 13 CTGCTGAAGAGATATGCGCATCAAGATGATGAGAGGCAACCTGTAACGGCCACTAC 72  
DB 22 CTGTTCAACGGGGTGTGCTCCATCTGTGAGCTGAGCGGCGACGTAAACGGCCACAAG 81  
QY 73 TTCAAGTCCGAGGGCGAGGGCGAAGCAACCTTCGCGGCGACCCAGAGCATGAGATC 132  
DB 82 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCGACCTAAGCGCAAGCTGAGCTTC 141  
QY 133 CACGTGACCGAGGGCGCGCCCTGCGCTTCGCTTGCATCCTGCGCCCTGCTGCGAG 192  
DB 142 ATCTGACACCA---CGGCAAGCTGCGCCCTGCGCTTGCATCCTGAGCACTTCGCGC 198  
QY 193 TACGCGACGAGACCTTCGTGACACCAACCGCGAGAT-----CCCCGACTTCTTCAAG 246  
DB 199 TACGCGCTGAAGTGTTCGCGCGCTACCGCGACCATGAGCAGCAGCACTTCTTCAAG 258  
QY 247 CAGAGCTTCCCGAGGGCTTCACTGAGAGAGAAACCACTTACGAGAGCGCGCGCATC 306  
DB 259 TCCGCCATGCCCCGAGGGCTACGTCCAGAGCGCACCATCTTCAAGAGCAGCGGCAAC 318  
QY 307 CTGACCGCCACCAAGACACCAAGCTGAGGCAACTGCTGATCTTCAAGGTGAAGTG 366  
DB 319 TACAAGACCCGCGCGAGGTGAAGTTCAGAGGCGACACCTGTGAACCGCATTCGAGCTG 378  
QY 367 CACGGCAACCACTTCCCGCGAGGGCGCGCGTGAAGAAACAAGAGCGCGCGCTGGAG 426  
DB 379 AAGGGCATGACTTCAAGAGAGGCGCAACATCTGGGGCACAAGCTGAGTACAACCTAC 438  
QY 427 CCCAGCACCGAGGTGTGTACCCCGAGAAACGGCGTGTGCGGGCGGAGCGTGA 481  
DB 439 AACAGCCACACGCTTATATATGCGCGACAGCAAGAAACGGCATCAAGGTGA 493

RESULT 11

US-09-602-641-14  
; Sequence 14, Application US/09602641  
; Patent No. 6608189  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Miyawaki, Atsushi  
; APPLICANT: Llopis, Juan  
; APPLICANT: Wachter, Rebekka M.  
; APPLICANT: Remington, S. James  
; TITLE OF INVENTION: FLUORESCENT PROTEIN SENSORS FOR  
; FILE REFERENCE: 07257/071001  
; CURRENT APPLICATION NUMBER: US/09/602,641  
; PRIOR FILING DATE: 2000-06-22  
; PRIOR APPLICATION NUMBER: 09/172,063  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 14  
; LENGTH: 720  
; TYPE: DNA  
; ORGANISM: Aequorea victoria  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)---(0)  
; OTHER INFORMATION: EYFP-V68L/Q69K  
US-09-602-641-14

Query Match 16.1%; Score 110.6; DB 4; Length 720;  
Best Local Similarity 55.2%; Pred. No. 2.8e-14;  
Matches 262; Conservative 0; Mismatches 204; Indels 9; Gaps 2;  
QY 13 CTGCTGAAGAGATATGCGCATCAAGATGATGAGAGGCAACCTGTAACGGCCACTAC 72

DB 22 CTGTTCAACGGGGTGTGCTCCATCTGTGAGCTGAGCGGCGACGTAAACGGCCACAAG 81  
QY 73 TTCAAGTCCGAGGGCGAGGGCGAAGCAACCTTCGCGGCGACCCAGAGCATGAGATC 132  
DB 82 TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCGACCTAAGCGCAAGCTTGAAGTTC 141  
QY 133 CACGTGACCGAGGGCGCGCCCTGCGCTTCGCTTGCATCCTGCGCCCTGCTGCGAG 192  
DB 142 ATCTGACACCA---CGGCAAGCTGCGCCCTGCGCTTGCATCCTGAGCACTTCGCGC 198  
QY 193 TACGCGACGAGACCTTCGTGACACCAACCGCGAGAT-----CCCCGACTTCTTCAAG 246  
DB 199 TACGCGCTGAAGTGTTCGCGCGCTACCGCGACCATGAGCAGCAGCACTTCTTCAAG 258  
QY 247 CAGAGCTTCCCGAGGGCTTCACTGAGAGAGAAACCACTTACGAGAGCGCGCGCATC 306  
DB 259 TCCGCCATGCCCCGAGGGCTACGTCCAGAGCGCACCATCTTCAAGAGCAGCGGCAAC 318  
QY 307 CTGACCGCCACCAAGACACCAAGCTGAGGCAACTGCTGATCTTCAAGGTGAAGTG 366  
DB 319 TACAAGACCCGCGCGAGGTGAAGTTCAGAGGCGACACCTGTGAACCGCATTCGAGCTG 378  
QY 367 CACGGCAACCACTTCCCGCGAGGGCGCGCGTGAAGAAACAAGAGCGCGCGCTGGAG 426  
DB 379 AAGGGCATGACTTCAAGAGAGGCGCAACATCTGGGGCACAAGCTGAGTACAACCTAC 438  
QY 427 CCCAGCACCGAGGTGTGTACCCCGAGAAACGGCGTGTGCGGGCGGAGCGTGA 481  
DB 439 AACAGCCACACGCTTATATATGCGCGACAGCAAGAAACGGCATCAAGGTGA 493

RESULT 12

US-09-459-956-2  
; Sequence 2, Application US/09459956  
; Patent No. 6342379  
; GENERAL INFORMATION:  
; APPLICANT: Tsien, Roger Y.  
; APPLICANT: Gonzalez, III, Jesus E.  
; TITLE OF INVENTION: DETECTION OF TRANSMEMBRANE POTENTIALS BY  
; FILE REFERENCE: REGEN1290-4  
; CURRENT APPLICATION NUMBER: US/09/459,956  
; PRIOR FILING DATE: 1999-12-13  
; PRIOR APPLICATION NUMBER: 08/765,860  
; PRIOR FILING DATE: 1999-05-08  
; PRIOR APPLICATION NUMBER: 08/481,977  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: PCT/US96/09652  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 690  
; TYPE: DNA  
; ORGANISM: Anemonia majano  
US-09-459-956-2

Query Match 16.0%; Score 109.8; DB 4; Length 690;  
Best Local Similarity 57.0%; Pred. No. 4e-14;  
Matches 223; Conservative 0; Mismatches 162; Indels 6; Gaps 1;

QY 44 ACATGAGAGGACCGGTGAACGGCCACTACTTCAAGTGCAGGGCGAGGCGCAAC 103  
DB 50 ATATGATGCTGTGTCAATGGGCACTTACCTTACCGTCAAGGTGAAGGCAACGGGAAGC 109  
QY 104 CTTTGGCGGACCCAGAGCATGAG-----AATCCAGTGAACCGAGGGCGCGCCCTGC 157  
DB 110 CATACGAAGGAGCAGACTTGAAGTCAAGTCAACCATGCGCAACGGTGGCCCTTG 169  
QY 158 CTTTGGCTTGAATCTGCGCGCGCGCTGCGAGTACGCGACGAGCACTTCTGTGACC 217  
DB 170 CATCTCTTGAATATCTATCAAGTGTTCAAATATGAATGATGATGCTTACTGCGT 229



[illegible]

```

RESULT 13
US-09-513-783A-43
; Sequence 43. Application US/09513783A
; Patent No. 6416959
; GENERAL INFORMATION:
; APPLICANT: Giuliani, Kenneth A.
; APPLICANT: Kapur, Ravl
; TITLE OF INVENTION: A System for Cell Based Screening
; FILE REFERENCE: 97-022-L1
; CURRENT APPLICATION NUMBER: US/09/513,783A
; CURRENT FILING DATE: 2000-02-25
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(717)
; OTHER INFORMATION: Description of Artificial Sequence: EYFP
US-09-513-783A-43

```

Query Match	15.9%	Score 109;	DB 4;	Length 717;
Best Local Similarity	54.9%	Pred. No. 5.9e-14;		
Matches 261; Conservative	0;	Mismatches 205;	Indels 9;	Gaps 2;

[illegible]

QY            427 CCCAGCACCGAGGTGTTACCCCGAGAACGGCTGCTGTGC CGGCCGGAACGTGA        481  
                |||||     |     |     |     |     |     |     |     |  
Db            439 AACAGCCACAAGTCTATATCATGGCCGACAGCAGAAAGAACGGCATCAAGGTGA        493

```

RESULT 14
; US-08-911-825-3
; Sequence 3, Application US/08911825
; Patent No. 6054321
; GENERAL INFORMATION:
; APPLICANT: Tsien et al., Roger Y. et al.
; TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/911,825
; FILING DATE: 15-AUG-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07257/056001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..720
;
US-08-911-825-3

```

Query Match	15.9%;	Score 109;	DB 3;	Length 720;
Best Local Similarity	54.9%;	Pred. No. 5.9e-14;		
Matches 261; Conservative	0;	Mismatches 205;	Indels 9;	Gaps 2;

Oy		13	CTGCTGAAGGAGAGTATGCCCATCAAGATGTACATGGAGGGCACCGTGAA CGGCCACTAC	72
Db		22	CTGTTCA CCGGGGTGTGCCCATCTGTGCAGCTGACCGCGCACTTAACGGCCACAAG	81
Oy		73	TTC AAGTGCAGGGCGAGGGCGA C GGCAACCCCTTCGCCGCGACCCAGA GATGAGAATC	132
Db		82	TTCAGCGTGTCCGGCGAGGGCGAGGGCGATTGCCACCCTAACGGCAAGCTGACCCTGAAGTTC	141
Oy		133	CACGTGACCGAGGGCGCCCCCTGCCCCTTCGCCCTTGACATCTGCGCCCCCTGCTGCGAG	192
Db		142	ATCTGCACCAC--CGGCAAGCTGCCCCGTGCCCTTGCCCAACCTCTGACCACTTCGGCC	198
Oy		193	TACGGCAGCAGACCTTCGTGCA CCA CACCGCCGAGAT-----CCCGACTTCTTCAAG	246
Db		199	TACGGCGTGCA GTGCTTCGCCCGCTACCCCGACCA CATGAAGCAGAGCACTTCTTCAAG	258
Oy		247	CAGAGCTTCCCCGAGGGCTTCACTTGGGAGAGAA CCA CCACTTACAGAGCAGCGCGGCATC	306
Db		259	TCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAAC	318
Oy		307	CTGACCGCCCA CCA GCA CCA CAGCCTGGAGGGCAACTGCCTGATCTACAAGGTGAAGGTG	366

Db 319 TACAAGACCCGCGCGAGGTGAAGTTCGAGGGCGACACCTGTGTGAACCCGCATCGAGCTG 378  
QY 367 CACGGCAACCACTTCCCCGCGACGCGCCCGTGATGAAGAACAGAGCGGCGCTGGAG 426  
Db 379 AAGGCATCGACTTCAAGGACGACGCAACATCCTGGGGCAAGAGCTGAGTACAACCTAC 438  
QY 427 CCCAGCAGGAGTGTGTATCCCCAGAACGGCGTGTGTGCGCGCGGAAGTGA 481  
Db 439 AACAGCCACAACGTCTATATCATGCGCGACAGCAAGAAAGGATCAAGGTGA 493

RESULT 15

US-08-974-737-3  
; Sequence 3, Application US/08974737  
; Patent No. 6077707

GENERAL INFORMATION:

APPLICANT: Tsien et al., Roger Y. et al.  
TITLE OF INVENTION: LONG WAVELENGTH MUTANT FLUORESCENT  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: USA  
ZIP: 92037

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,737  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/911,825  
FILING DATE: 15-AUG-1997

ATTORNEY/AGENT INFORMATION:  
NAME: Hall, Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07257/056001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 720 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..720  
US-08-974-737-3

Query Match 15.9%; Score 109; DB 3; Length 720;  
Best Local Similarity 54.9%; Pred. No. 5.9e-14;  
Matches 261; Conservative 0; Mismatches 205; Indels 9; Gaps 2;

QY 13 CTGCTGAAGAGATATGCGCATCAAGATGTACATGAGGGCACCGTGAACGGCCACTAC 72  
Db 22 CTGTTCAACGGGTGTGCTCCATCTGTGTGAGCTGACGCGACGTAAACGGCCACAAG 81  
QY 73 TTCAAGTCGAGGGCGAGGGCGACGGCAACCCCTTGGCGGCACCCAGAGCATGAGAATC 132  
Db 82 TTCAAGCTGTCCGGCGAGGGCGAGGGCGATGCCACTTACGGCAAGCTGACCTGAAGTTT 141  
QY 133 CACGTACCGAGGGCGCGCCCTGCTTGGCTTGCACATCTGGCCCCCTGCTGCGAG 192  
Db 142 ATCTGACACAC--CGGCAAGCTGCCCTGCGCCCAACCTGTGACCACTTGGGC 198

QY 193 TACGGCAGAGGACCTTCTGTGCACCAACGACCCCGAGAT-----CCCCACTTCTTCAAG 246  
Db 199 TACGGCGTCAAGTGTCTGCGCCCGCTACCCCGACACATGAAGCAGAGACTTCTTCAAG 258  
QY 247 CAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCAACCTACGAGAGCGGCGCATC 306  
Db 259 TCCGCCATGCCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGAGCAGCGCAAC 318  
QY 307 CTGACCGCCCAACGACACAGCCTGAGGGCAACTGCTGATCTTACAAGGTGAAGGTG 366  
Db 319 TACAAGACCCGCGCGAGGTGAAGTTGAGGGCGACACCTGTGTGAACCGCATCGAGCTG 378  
QY 367 CACGGCAACCACTTCCCCGCGACGCGCCCGTGATGAAGAACAGAGCGGCGCTGGAG 426  
Db 379 AAGGCATCGACTTCAAGGACGACGCAACATCCTGGGGCAAGAGTGAACAACCTAC 438  
QY 427 CCCAGCAGGAGTGTGTATCCCCAGAACGGCGTGTGTGCGCGCGGAAGTGA 481  
Db 439 AACAGCCACAACGTCTATATCATGCGCGACAGCAAGAAAGGATCAAGGTGA 493

Search completed: July 15, 2004, 14:06:51  
Job time : 85 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2004, 11:15:43 ; Search time 2886 Seconds  
(without alignments)  
7108.565 Million cell updates/sec

Title: US-09-976-673-11

Perfect score: 687

Sequence: 1 atggtgagcgagcgtgctgaa.....tgcccgagaagccaactga 687

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vr1:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	99.4	14.5	550	13 BQ551001	BQ551001 H4006G01-AI389106 GH20192.5
2	71	10.3	645	9 AI389106	AI389106 GH20192.5
c 3	68.6	10.0	869	14 CK159167	CK159167 FGAS04056
4	67.6	9.8	639	9 AJ615477	AJ615477 AJ615477

5	67.4	9.8	675	12 BI382638	BI382638 BFLG2_000
6	66.8	9.7	944	29 CC649516	CC649516 OGLC16TH
7	66.4	9.7	700	14 CF864358	CF864358 psZS009XP
8	65.8	9.6	624	28 B2422085	B2422085 1d49d07.b
9	65.8	9.6	982	28 CC424389	CC424389 PUHNS3TB
c 10	65.2	9.5	794	14 CB651092	CB651092 OSJNEb15O
c 11	64.8	9.4	697	13 BU626888	BU626888 UI-H-FT0-
12	64.6	9.4	821	14 CK199961	CK199961 FGAS00846
13	64.2	9.3	613	12 BI386976	BI386976 BFL26_001
14	64.2	9.3	643	12 BI386699	BI386699 BFL26_001
c 15	64	9.3	681	29 CNS02EOD	AL193990 Tetradon
c 16	64	9.3	688	29 CG038305	CG038305 PUFNQ30TD
17	63.8	9.3	411	10 BF201975	BF201975 WHE1759-1
18	63.8	9.3	600	10 BE417750	BE417750 MUG024.B0
19	63.8	9.3	772	13 BQ752733	BQ752733 WHE4118_E
20	63.6	9.3	872	28 CC413363	CC413363 PUHPU48TD
21	63.6	9.3	956	29 CG365299	CG365299 OGWFT62TV
c 22	63.4	9.2	925	29 CNS0091P	AL053013 Drosophil
23	62.8	9.1	1006	29 CC719047	CC719047 OGULN81TV
24	62.6	9.1	599	9 AI832082	AI832082 tcl2c04.x
25	62.6	9.1	689	12 BI380055	BI380055 BFLG1_001
26	62.4	9.1	1036	14 CK162903	CK162903 FGAS01551
27	62.2	9.1	638	12 BJ269933	BJ269933 BJ269933
28	62.2	9.1	742	14 CB649099	CB649099 OSJNEb12N
29	62.2	9.1	780	14 CF554490	CF554490 Ggw74 Inf
30	62	9.0	786	14 CB646357	CB646357 OSJNEb08M
31	62	9.0	982	13 BX415111	BX415111 BX415111
32	61.8	9.0	557	12 BJ233355	BJ233355 BJ233355
33	61.6	9.0	590	13 CA030137	CA030137 HX06D04r
34	61.4	8.9	540	12 BM135283	BM135283 WHE0498_B
35	61.4	8.9	635	14 CA254345	CA254345 SCBFFL411
36	61.2	8.9	724	14 CF871694	CF871694 tRic028xf
37	61.2	8.9	739	14 CB645917	CB645917 OSJNEb07P
38	61.2	8.9	777	14 CB901927	CB901927 tRic028xf
39	61	8.9	571	12 BI387406	BI387406 BFL26_002
40	61	8.9	661	12 BI385436	BI385436 BFL26_000
41	61	8.9	687	12 BI387917	BI387917 BFL26_002
42	61	8.9	502	14 BI135401	CB657967 OSJNEc13K
43	60.8	8.9	502	12 BI135401	BI135401 6HRM115_6
c 44	60.6	8.8	559	14 CF847678	CF847678 pGMA002xB
c 45	60.6	8.8	759	28 BZ577913	BZ577913 msh2_5633

ALIGNMENTS

RESULT 1  
BQ551001  
LOCUS  
DEFINITION H4006G01-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone  
H4006G01 5', mRNA sequence.  
ACCESSION BQ551001  
VERSION BQ551001  
KEYWORDS  
SOURCE EST.  
ORGANISM Mus musculus (house mouse)  
MUS musculus  
REFERENCE  
AUTHORS VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,  
Martin,P.R., Stagg,C.A., Bassey,U., Alba,K., Hamatani,T.,  
Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.  
TITLE Assembly, verification, and initial annotation of NIA 7.4K mouse  
cDNA clone set  
JOURNAL Genome Res. 12 (12), 1999-2003 (2002)  
MEDLINE 22354164  
PUBMED 12466305  
COMMENT Other ESTs: H4006G01-3  
Contact: Yong Qian  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA  
Email: cdna@lgsun.grc.nia.nih.gov





REFERENCE 1 (bases 1 to 869)  
AUTHORS Allard,F., Crosby,W.L., Danyluk,J., Budes,F., Frick,M., Gaudet,D., Genswein,B., Graf,R., Gulick,P., Hrycan,L.D., Laroche,A., Links,M.G., McCarthy,E.L., Monroy,A., Muzak,I., Nilsson,D., Peniket,C., Roach,J.L. and Sarhan,F.  
TITLE Functional Genomics of Abiotic Stress In Wheat and Canola Crops  
JOURNAL Unpublished (2003)  
COMMENT Contact: Wm L Crosby  
Bioinformatics  
University of Saskatchewan, Department of Computer Science  
1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
Saskatchewan, S7N 5A9, Canada  
Tel: 306 966 1769  
Fax: 306 966 2033  
Email: fgas\_estes@cs.usask.ca  
This sequence is the direct result of the Base calling software  
phred (default parameters). It is the raw base calls. To aid in the  
identification of the high quality insert the software Lucy  
(default parameters) has been run on this sequence. Lucy identified  
the region [128,636].  
Plate: Talc537 row: N column: 23.  
Location/Qualifiers  
1..869

FEATURES

source

/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/cultivar="wheat line PI 178383"  
/db\_xref="taxon:4565"  
/lab\_host="DHS alpha"  
/clone\_lib="Triticum aestivum FGAS: Talc5"  
/note="Organ: Crown; Vector: pGEM-T; SSH (suppression  
subtractive hybridization) cDNA library from genotype  
PI178383 cold hardened at 2 C for 21 days and 49 days  
(equal amount of cDNA pooled together before subtraction,  
tester) and subtracted against genotype Norstar cold  
hardened at 2 C for 1 day (24 H)(driver). Modified Smart  
cDNA (Clontech)priming and non-directional cloning"

ORIGIN

Query Match 10.0%; Score 68.6; DB 14; Length 869;  
Best Local Similarity 44.2%; Pred. No. 2.3;  
Matches 284; Conservative 0; Mismatches 359; Indels 0; Gaps 0;

QY 24 GAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCACTACTTCAAGTGGCA 83  
DB 807 GAGGATCACCAACGAGAGCAACCAACCAACCAACCAACGACCAACGACGACCAACGCA 748  
QY 84 GGGCGAGGGCGACGCAACCCCTTCGCCGGCACCCAGACGATGAGATCCAGTACGACCGA 143  
DB 747 CACAACGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 688  
QY 144 GGGGCCCCCCTGCGCTTCGACATCCTGGCCCCCTGCTGCGAGTACGGCAGCAG 203  
DB 687 CAACCAACAACAACGACCAACCAACCAACCAACCAACAACAACAACAACAACA 628  
QY 204 GACCTTGTGTCACCAACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCGGAGG 263  
DB 627 CAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 568  
QY 264 CTTCACCTGGAGAGAAACCAACCACTAGAGAGCGCGGCAATCCTGACCGCCACCAAGA 323  
DB 567 CAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 508  
QY 324 CACCAACCTGGAGGCAACTGCTGATCTACAAGGTGAAGTGCAAGGACCAACTTCCC 383  
DB 507 CAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 448  
QY 384 CGCGCAGCGCCCCGTGATGAAGAACAGAGCGCGGCTGGAGCCCAAGCAGCAGGTGT 443  
DB 447 CAACCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 388  
QY 444 GTACCCCGAAGACGGCGTGTGTGCGGCGCGGAACGTGATGGCCCTGAAGGTGGCGACCG 503  
DB 387 CAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 328

QY 504 GCACCTGATCTGCCACCACTACACCAAGTACCGGACGAAGAGCCGTGCGCCCTGAC 563  
DB 327 CAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 268  
QY 564 CATGCCCGGCTTCCACTTCACCGACATCCGGCTCCAGATGCTGCGGAAGAAGACGA 623  
DB 267 CAACCAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 208  
QY 624 GTACTTCAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCGAC 666  
DB 207 CGACAACGACGTCACGATTAACAACGACACAACAACGACACAACAAC 165

RESULT 4  
AJ615477 639 bp mRNA linear EST 11-DEC-2003  
LOCUS AJ615477 Triticum turgidum subsp. durum etiolated seedling 20 day  
DEFINITION Triticum turgidum subsp. durum cDNA clone 11702R, mRNA sequence.  
ACCESSION AJ615477  
VERSION AJ615477.1 GI:39732404  
KEYWORDS EST.  
SOURCE Triticum turgidum subsp. durum (durum wheat)  
ORGANISM Triticum turgidum subsp. durum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.  
1 (bases 1 to 639)  
Cifarelli,R.A., D'Onofrio,O. and Lauria,G.  
Expressed Sequences Tags (ESTs) library from totipotent cDNA of  
durum wheat  
Unpublished (2003)  
Contact: Cifarelli RA  
Biotechnology  
Metapontum Agrobios  
S.S. Jonica 106 Km 448.2, 75010 Metaponto (MT), Italy.  
Location/Qualifiers  
1..639

REFERENCE 1 (bases 1 to 639)  
AUTHORS Cifarelli,R.A., D'Onofrio,O. and Lauria,G.  
TITLE Expressed Sequences Tags (ESTs) library from totipotent cDNA of  
durum wheat  
JOURNAL Unpublished (2003)  
COMMENT Contact: Cifarelli RA  
Biotechnology  
Metapontum Agrobios  
S.S. Jonica 106 Km 448.2, 75010 Metaponto (MT), Italy.  
Location/Qualifiers  
1..639

FEATURES  
source  
/organism="Triticum turgidum subsp. durum"  
/mol\_type="mRNA"  
/cultivar="Ofanto"  
/sub\_species="durum"  
/db\_xref="taxon:4567"  
/clone="11702R"  
/tissue\_type="etiolated seedling"  
/dev\_stage="20 day"  
/clone\_lib="Triticum turgidum subsp. durum etiolated  
seedling 20 day"

ORIGIN

Query Match 9.8%; Score 67.6; DB 9; Length 639;  
Best Local Similarity 52.9%; Pred. No. 3;  
Matches 145; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 58 GTGAACGGCCACTACTTCAAGTCCGAGGGCGAGGGCAACCCCTTCGCGCGCAAC 117  
DB 294 GTGCTGGCGCTCAGCAAGAGTGCTGTGCTGTGCTCAAGGACCGGACGAGCCACC 353  
QY 118 CAGAGCATGAGATCCACGTGACCGAGGGCGCCCGCTGCTGCTTGCCTTGACATCCTG 177  
DB 354 CTCGGATCAAGTTCAACGTCAACCCGCGCATGAACCTCCCTCGCTGCAACATCCCG 413  
QY 178 GCCCCCTGCTGCGATGAGGACGACGACCTTGTGACCAACCGCGAGATCCCGAC 237  
DB 414 GCCACCTTCTCCGACTGCCCAAGATCTCAACATGTGCGCGGACTCCAAGAGGCCGAG 473  
QY 238 TTCTTCAAGCAGAGCTTCCCGAGGGGCTTCACTTGGAGAGAACAACCACTACGAGAC 297  
DB 474 ATCTTCAAGCAGTACGGGATCGAGCAGCAGGCGCAAGAACGCCACCGCGGCGAGCC 533  
QY 298 GGGCGCATCTGACCGGCCCAAGGACCAAGACCGC 331  
DB 534 GCCGTACCGGTACTCTCCGGCGGGAAGAGCGCC 567

RESULT 5  
BI382638  
LOCUS  
DEFINITION  
BI382638 675 bp mRNA linear EST 26-AUG-2003  
BFLG2\_000815 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG  
or MPMGp498) Branchiostoma floridae cDNA clone MPMGp498F0258 5',  
mRNA sequence.  
ACCESSION  
BI382638  
VERSION  
BI382638.1 GI:30917704  
KEYWORDS  
EST.  
SOURCE  
Branchiostoma floridae (Florida lancelet)  
ORGANISM  
Branchiostoma floridae  
Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
Branchiostoma.  
REFERENCE  
1 (bases 1 to 675)  
Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J.,  
Herwig,R., Vingron,M. and Lehrach,H.  
New evidence for genome-wide duplications at the origin of  
vertebrates using an amphioxus gene set and completed animal  
genomes  
JOURNAL  
Genome Res. 13 (6A), 1056-1066 (2003)  
MEDLINE  
22683279  
PUBMED  
12799346  
COMMENT  
Contact: Panopoulou G  
laboratory 145, dept.Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inestr.63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoul@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONFP) to reduce sequencing redundancy. According to the ONFP  
procedure, clones giving the same hybridisation pattern with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One  
clone per cluster is selected for sequencing. The size of each  
cluster is an indicator of the frequency of a transcript in the  
analysed library. The cluster size as well the coordinates of the  
rest of the clones assigned to the same fingerprint cluster as the  
clone from which the above EST is generated is available at the  
amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>  
Clones and filters are distributed via the Resource Center/Primary  
Database of the German Genome Project (<http://www.rzpd.de>)  
PCR Primers  
FORWARD: 5' CCCACGGCTTACACTTATGCTTCGGGCTCG 3' (M13RSP)  
BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)  
Insert Length: 1 Std Error: 0.00  
Seq primer: 5'-CCGGTCCGAATTCCTCCGGT-3' pSPORT3/86  
High quality sequence stop: 675.  
FEATURES  
source  
Location/Qualifiers  
1..675  
/organism="Branchiostoma floridae"  
/mol\_type="mRNA"  
/db\_xref="taxon:7739"  
/clone="MPMGp498F0258"  
/tissue\_type="whole embryo"  
/dev\_stage="5-6 hrs (gastrula stage)"  
/lab\_host="E.coli, XLI blue"  
/clone\_lib="Amphioxus 5-6 hrs cDNA library (Name  
convention: BFLG or MPMGp498)"  
/note="Vector: pSPORT1; Site\_1: SalI, KpnI, EcoRI (5');  
Site\_2: NotI, BamHI, HindIII (3'); OligodT primed and  
directionally cloned in pSPORT1 vector using a NotI  
(5'-TGACTAGTCTAGATCGGAGCGGCGCCC (T)15-3' and a SalI 5'-  
TCGACCCACGCGTCCG-3' adapters (Gibco BRL)."

ORIGIN  
Query Match 9.8%; Score 67.4; DB 12; Length 675;  
Best Local Similarity 48.1%; Pred. No. 3.2;  
Matches 191; Conservative 0; Mismatches 206; Indels 0; Gaps 0;  
QY 36 CAAGATGTACATGAGGCGACCGTGAACGCCCACTACTTCAAGTGCAGAGGCGAGGCGGA 95  
|||||

Db 106 CGACATCCACCTTCACGGCTCCATCAACGGCCAGAGTTCCATGTGGGGAGGAAA 165  
QY 96 CGGCAACCCCTTCGCCCGGACCCAGAGCATGAGAAATCCACCTGACCGAGGGCGCCCCCT 155  
Db 166 AGGCGACCCCGAAGCGCCGCTCGCTGTGACACACAGGAAATCCACCAAGGCTCCCTGAA 225  
QY 156 GGCCTTCGCCCTTCGACATCTGAGCCCCCTGCTGCGAGTACGGCAGCAGACCTTCGTGA 215  
Db 226 GTTCTCCCCCTTACTTGATGATCCCCCACCCTCGGGTACGGGTACTACCACTACCTCCCTA 285  
QY 216 CCACACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCCGAGGCTTCACTGGGA 275  
Db 286 CCGGACGACCCCTCGCCTTTCCAGGTCTTCATGTTGGAAGATCGGGTATGCAGTCTA 345  
QY 276 GAGAACCAACCACTACGAGACGGCGGCAATCCTGACCCGCCACAGACACCAAGCCTTGA 335  
Db 346 CCGCGTGTTCGACTTTGAAGACGAGGCAAGCTGACTACCGAGTTTAAGTACTCTTACGA 405  
QY 336 GGGCAACTGCTTGATCTPACAAAGTGAAAGTGACACGGCAACCACTTCCCCCGGACGGGCC 395  
Db 406 GGGTTCCTATATCAAGGCCGACATGAACTGATGGAAAGCGGTTCCCTGACGACGGGCC 465  
QY 396 CGTGATGAAGAACAAAGACGGCGGCTGGAGCCACG 432  
Db 466 AGTCATGACCAAGCAGATTGTTCGACCAGACGCGCTGC 502

RESULT 6  
CC649516  
LOCUS  
DEFINITION  
CC649516 944 bp DNA linear GSS 19-JUN-2003  
genomic survey sequence.  
ACCESSION  
CC649516  
VERSION  
CC649516.1 GI:32052184  
KEYWORDS  
GSS.  
SOURCE  
Zea mays  
ORGANISM  
Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
1 (bases 1 to 944)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Uterback,T.,  
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,  
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.  
Consortium for Maize Genomics  
Unpublished (2002)  
OTHER\_GSSs: OGLCL16TV  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.  
FEATURES  
source  
Location/Qualifiers  
1..944  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMBMA0322D08"  
/clone\_lib="ZM\_0.7-1.5\_KB"  
/note="Vector: pBCSK-; Site\_1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

ORIGIN  
Query Match 9.7%; Score 66.8; DB 29; Length 944;  
Best Local Similarity 45.7%; Pred. No. 4.2;  
Matches 273; Conservative 0; Mismatches 322; Indels 3; Gaps 1;  
QY 30 GCGCATCAAGATGATGAGAGGCGACCGTGAACGGCCACTACTTCAAGTGCAGAGGCGGA 89  
|||||

Db 129 GCGCCTCGCCCTGTCTACGCGCGCGCGCGCTGCGGCTCGTGCACCCCTCGGCGCGCAC 188  
QY 90 GGGCGCAGCGCAACCCCTTCGCCCGGCACCAGAGCATGAGATCCACGTGACCGAGGGCGC 149  
Db 189 CGGCGTCGCCAACCGCGGACTGTCTACTTCAACGGCGCGCTCTCGCCATGTCCGAGGA 248  
QY 150 CCCCCTGCCCTTCGCTTCGACATCTGCGCCCTGCTGCGAGTACGGCAGACGACCTT 209  
Db 249 CGACCTGCGGTACCACTCCGCGTCGCGGAGCAGACGGCACTTCGAGACCGTCCGCGCTA 308  
QY 210 CGTGCACCAACCGCGGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCGAGGCTTCAC 269  
Db 309 CGACTTCGACGGCCAGCTCGGCTGCCCCATGATCGCGCACCCCAAGCTGGAACCGGCGCAC 368  
QY 270 CTGGGAGAGAACCACCACTACGAGAGCGCGGCATCCTGACCGCCCAAGCAGACAGCAG 329  
Db 369 CGGGGAGCTGCACGCGCTCAGCTACGAGGTGTCAGGAGGCGCTTACCTCAAGTACTTCTA 428  
QY 330 CCTGGAGGGCAACTGCTGA--TCTACAAGGTGAAGGTGACGGCACTTCCCGC 386  
Db 429 CTTCAAGCGCCGACGGCACTCAAGTCCGACGAGTGAGATCCGCTGGCCCCACCATCAT 488  
QY 387 CGACGCGCCCGTGATGAAGACAAGCGCGCGCTGGAGGCCAGCACCGAGGTGTGTA 446  
Db 489 GATCCACGACTTCGCCCATCACCGAAGAACTTGTCGTGTGCCCGACCAACCAAGTGTGTT 548  
QY 447 CCCCAGAACGGCGTGTGTGCGCGCGGAACGTGATGGCCCTGAAGGTGGCGACCGCA 506  
Db 549 CAAGCTGCAGAGAGATGCTGCGCGCGGGTCCGCCGTGCTGGAACAGGAGAAGACGTC 608  
QY 507 CCTGATCTGCCACCACTACACGACTACCGGAGCAAGAGCGCGTGCGCCCTGACCAT 566  
Db 609 GCGCTTCGCGCTCCTCCCGAAGCGCGCGCGGACGCGCTCGAGATGGCGGTGAGCGT 668  
QY 567 GCCCGGCTTCCACTTCACCGCAATCCGGCTCCAGATGCTCGCGAAGAAGAAGACGAG 624  
Db 669 GCCGGACTGCTTCTGCTTCACCTGTGGAACGCGTGGAGGACGAGCGGCGCGAG 726

RESULT 7  
CF864358 700 bp mRNA linear EST 31-OCT-2003  
LOCUS CF864358  
DEFINITION pez5009xP17f USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_szs Phytophthora sojae cDNA clone sz5009P17 5, mRNA sequence.  
ACCESSION CF864358  
VERSION CF864358.1 GI:38118984  
KEYWORDS EST.  
SOURCE Phytophthora sojae  
ORGANISM Phytophthora sojae  
Phytophthora  
Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae; Phytophthora.  
REFERENCE 1 (bases 1 to 700)  
AUTHORS Tyler,B.  
TITLE Tyler,B. Not Published  
JOURNAL Unpublished (2003)  
COMMENT Contact: Tyler B  
Tyler lab

VBI  
1880 Pratt Dr., Blacksburg, VA 24061, USA  
Tel: 540-231-7318  
Email: bmtyley@vt.edu

PCR Primers  
FORWARD: BK reverse primer  
BACKWARD: BK reverse primer  
Plate: 009 row: P column: 17  
Seq primer: BK reverse primer  
High quality sequence stop: 700.

FEATURES  
source  
1..700  
/organism="Phytophthora sojae"  
/mol\_type="mRNA"  
/db\_xref="taxon:67593"

/clone="sz5009P17"  
/tissue\_type="Zoospores"  
/cell\_line="P6497"  
/dev\_stage="Free swimming"  
/clone\_id="USDA-IPAFS:Expression of Phytophthora sojae genes during infection and propagation\_szs"  
/note="Vector: PBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

Query Match 9.7%; Score 66.4; DB 14; Length 700;  
Best Local Similarity 45.7%; Pred. No. 4.5;  
Matches 269; Conservative 0; Mismatches 316; Indels 3; Gaps 1;

QY 66 CCACTACTTCAAGTGCAGGGCGAGGGCGAGCGCAACCCCTTCGCCCGGCACCAGAGCAT 125  
Db 103 CCGCCACACCAAGAGCCACAGCGTGGATGGCATAGCCCGCGGCACTTCACCGT 162  
QY 126 GAGAATCCACGTGACCGAGGGCGCCCCCTTGCCCTTCGACATCCTGGCCCCCTG 185  
Db 163 GAG--CTCGTCGCCAGCTCGCCGCGGACCAAGCAAGTCCGCCACAGCGCCGCG 219  
QY 186 CTGCGAGTACGCGACAGACCTTCGTGACCAACCGCGGAGATCCCGACTTCTCAA 245  
Db 220 CTTCCACAGGGTCAATGGCTGACGAGGTGCTCCCAAGAGCGCATGACGTATCTA 279  
QY 246 GCAGAGCTTCCCGAGGGCTTCACTGGAGAGAACCACTACGAGGACGGCGCAT 305  
Db 280 GACTACTCTCATGAGCGCATCTCGAGTGCGCCCGCTCATCGCGTGTGTGCGGCTCCG 339  
QY 306 CTTGACCGCCCAAGAGACCAAGCCTTGAGGGCACTGCTGATCTTCAAGTGAAGT 365  
Db 340 CTTGGAGACTCAGCAACTGCTCTGTCCAACACGAGAACATCAAGTACGAGAGATCCC 399  
QY 366 GCACGGCACCAACTTCCCGCGGACGCGCCCGTGATGAAGAACAGCGCGGTGGA 425  
Db 400 GCAATTCCCGGCTCCACCGTGCAGGGCCACGCGGAGAGCTGCTTCGGCGACTGGA 459  
QY 426 GCCCAGCACCGAGGTGTGTACCCCCGAGAACGGCGTGTGCGCGCGGAACGTGATGCC 485  
Db 460 GGGCTTCCGCGTGTGTGTCATGCGCGCGCTTCCACTGCTACGAGGGCTACGCCATGCG 519  
QY 486 CTGAAGGTGGGCGGACCGGCACTGATCTGCCACCACTACACGACTACCGGAGCAAGA 545  
Db 520 CGAGACGGCGTGCCCATTCGCGTCATGTACTCTCTCGGCATCAAGTACTTGTGTGAC 579  
QY 546 GCGCGTGCGCGCCCTGACCATGCCGCGCTTCACCTTCAACGACATCCGGCTCCAGTGT 605  
Db 580 CAACGCGCGGGCGCGCTGAACCCGGACTTCAACGTGGGCGAGCTGATGATCTGAACGA 639  
QY 606 GCGGAAGAAGAAGCAGTACTTTCGAGCTGTACGAGGCCAGCGTGGC 653  
Db 640 CCACCTGAACGTGCCCGGCTGTGCGGTACAGCACCGGCTCATCGGGCC 687

RESULT 8  
BZ422085 624 bp DNA linear GSS 10-DEC-2002  
LOCUS BZ422085  
DEFINITION id49d07.b1 WGS-Sbicolorf (DH5a methyl filtered) Sorghum bicolor genomic clone id49d07 5', genomic survey sequence.  
ACCESSION BZ422085  
VERSION BZ422085.1 GI:26370029  
KEYWORDS GSS.  
SOURCE Sorghum bicolor (sorghum)  
ORGANISM Sorghum bicolor  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Sorghum.  
REFERENCE 1 (bases 1 to 624)  
AUTHORS Rabinowicz,P.D., O'Shaughnessy,A.L., Balijs,V., Dedhia,N., Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L., Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A.  
TITLE Genomic shotgun sequences from Sorghum bicolor (methyl-filtered) Unpublished (2002)

## COMMENT

Contact: W. Richard McCombie  
Lita Annenberg Hazen Genome Sequencing Center  
Cold Spring Harbor Laboratory  
PO Box 100, Cold Spring Harbor, NY 11724, USA  
Tel: 516 367 8884  
Fax: 516 367 8874  
Email: mcombie@cshl.org  
Plate: id49 row: d column: 07  
Seq primer: -21M13UnivFwd  
Class: Shotgun  
High quality sequence stop: 624.

## FEATURES

## source

1..624  
/organism="Sorghum bicolor"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:4558"  
/clone="id49d07"  
/lab\_host="DH5a"  
/clone\_lib="WGS-SbicolorF (DH5a methyl filtered)"  
/note="Site 1: Xba I; Site 2: Xba I; The vector was digested with Xba I and one nucleotide was added by fill in in the recessive 3' end. The genomic DNA was nebulized, end repaired, adaptor ligated and size fractionated using sephadex. The resulting fragments were between 0.8 and 3 kb and were cloned into the vector (.x/y reads in M13mp19, .b/g reads in pUC19). The same ligation was transformed into DH5a."

## ORIGIN

Query Match 9.6%; Score 65.8; DB 28; Length 624;  
Best Local Similarity 48.1%; Pred. No. 5.4;  
Matches 187; Conservative 0; Mismatches 202; Indels 0; Gaps 0;  
QY 79 TCGAGGGCGAGGGCGAGCAACCCCTTCGCCGCGCAGACCATGAGATCCACGTG 138  
Db 58 TTCACGTACGGTGTGGGACGACGACGTGGCGATCCACCGGTGAGCATCCGGCG 117  
QY 139 ACCGAGGGCCCCCTTGCCTTCGACATCCTGGCCCCCTGCTGGAGTACGGC 198  
Db 118 CCGCGAGAGCCCCCAACACGAGCGCGTGCATCCAGGGCTCTCCACGTCCGGCTC 177  
QY 199 AGCAGACCTTGTGACCAACCGCGGAGATCCCGACTTCTCAAGCAGAGCTTCCC 258  
Db 178 ACGACTCGGCGCGTGCACACCGGCGACGATGCTCTCCGTCGGCGCCGGCGCTCCGAC 237  
QY 259 GAGGCTTCACTTGGAGAGAACCACTTAAGAGAGCGCGCATCTGACCGCCAC 318  
Db 238 GTGTGTGTTCGGGGTGTGTGTGCGGCGCGGCGACGCGATCAGCGTCCGACGCTCGGC 297  
QY 319 CAGGACACCAAGCTTGGAGGCACTGCTGATCTTCAAGTGAAGTGCAAGGCAACCAAC 378  
Db 298 AGTACCGCGCGAGAGAGAGTCCGCGGCTGCGGCGCACTGACCTTCCGCGGC 357  
QY 379 TTCCCGCGCAGCGCGCGGATGAAGAACAGAGCGGCGCTGGAGCCACGACCGAG 438  
Db 358 ACCTCAACGCGCGTGCATCAAGAGTGGCGCGCGGCACTTGGCCACCGCGCTCAC 417  
QY 439 GTGTGTACCCCGAGAACGCGTGTGTG 467  
Db 418 GGCTGTCTTCGAGGACATCGTCATGAG 446

## RESULT 9

## CC424389

LOCUS CC424389 982 bp DNA linear GSS 19-MAY-2003  
DEFINITION PUHMK53TB\_ZM\_0.6\_1.0\_KB Zea mays genomic clone ZMMBTA491I09,  
genomic survey sequence.

## ACCESSION

## CC424389

## VERSION

## CC424389.1

## GI:30904479

## GSS.

## KEYWORDS

## Zea mays

## Zea mays

## ORGANISM

## Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
1 (bases 1 to 982)  
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and  
Bennetzen, J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other\_GSSs: PUHMK53TD  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

## source

1..982  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone\_lib="ZMMBTA491I09"  
/note="Vector: PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
cot selected genomic DNA library"

## ORIGIN

Query Match 9.6%; Score 65.8; DB 28; Length 982;  
Best Local Similarity 46.2%; Pred. No. 5.9;  
Matches 294; Conservative 0; Mismatches 337; Indels 6; Gaps 2;  
QY 28 ATGCGATCAAGATGATGATGAGGAGCCGTTGAACGGCCACTTCAAGTGCAGGGC 87  
Db 263 AAGGCGCTCAGTGCCTGCGCGCAGCTGATCAACCAACCGCTGCGCGCACTACAAGAGAC 322  
QY 88 GAGGCGAGCGCAACCCCTTGCCTGCGCGCAGAGCATGAGAAATCCACGTGACCGAGGGC 147  
Db 323 GGCCGCGGATCTACTGCTGCTTTCGAGGGCGGCAAGCGCGCAAGCCGCTTCACTGGGGC 382  
QY 148 GCGCCCTGCTTGCCTTGCATCTG--GCGCCCTGCTGCGAGTACGGCAGCAGAG 204  
Db 383 CCGCATGATCTGAGAGAGACGACGAGTACTCCAAAGGGCGGGGCGCGGACACG 442  
QY 205 ACCTTGTGACACACCGCGCGAGATCCCGACTTCTCAAGCAGAGCTTCCCGAGGGC 264  
Db 443 GGGGCGCACTTGCCTGCGCGCGCGCGCATGACCACTCAACCGCGCTGCAGCAGAGAG 502  
QY 265 TTCACTGGAGAGAACCACTTAAGAGAGCGCGCATCTGACCGCCACCAAGAGAC 324  
Db 503 CTCTCGACTGCTCAACTGGCTCAAGTCCGACTTGGCTTGAAGGCTGGCGCTCGAC 562  
QY 325 ACCAGCTGAGAGGCACTGCTGATCTAAGAGTGAAGTGCACGGCAACCACTTCCC 384  
Db 563 TTCGCAAGGCTACTCCGCGCGCGCGCAAGTGTACGTGACAGCAGCAGCGCCCGCAC 622  
QY 385 GCCGAGCGCGCGTGTGAAGAACAGAGCGCGGCTGGAGCCAGCACCGAGTGTG 444  
Db 623 TTGTGTGTGCGGAGATGAGAGC--TCCCTCACTAAGACGCAAGCGGAGCGCTCC 679  
QY 445 TACCCGAGAGAGCGCGTGTGCGCGCGGAACGTGATGGCCCTGAAGTGGCGAGCCGG 504  
Db 680 AGCAACCAAGAGCGCGAGAGAGCTGTCAACTGGCGGAGCGGCTGGCGCGCGCC 739  
QY 505 CACTGATCTGCGACCACTACACCAAGTACCGGAGCAAGAGCGCGCGCGCTGACC 564  
Db 740 GCCGCGCGCTTGCATTCACCAACCAAGGCGGTGCTGCAAGGCGCGCGCTTCAAGGCGAGCTG 799  
QY 565 ATGCCCGGCTTCACTTCAACGCAATCCGGCTCCAGATGCTGCGGAGAGAGAGAGAG 624  
Db 800 TGCGCATGAAGAGCGCAACCGGAGAGCGCGCGGATGATCGGCTGCGCGAGAG 859



Oy	625	TACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGTACA	661
Db	860	GCCGTCACGTTCCGTGCACACCACGACCCGGCTCCA	896
RESULT 10			
LOCUS CB651092/c		794 bp	mRNA linear EST 08-APR-2003
DEFINITION OSJNEB15018.r OSJNEB Oryza sativa (japonica cultivar-group) cDNA clone OSJNEB15018 3', mRNA sequence.			
ACCESSION CB651092			
VERSION CB651092.1 GI:29646085			
SOURCE EST.			
ORGANISM Oryza sativa (japonica cultivar-group)			
Oryza sativa (japonica cultivar-group)			
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;			
Ehrhartoidae; Oryzeae; Oryza.			
1 (bases 1 to 794)			
Jantaauriyarat,C., Lu,G., Gowda,M., Hatfield,J., Zhou,B., Mazur,E., Kudrna,D., Dean,R., Soderlund,C., Wing,R. and Wang,G.			
large-scale identification of ESTs involved in the interaction between rice and Magnaporthe grisea			
Unpublished (2003)			
REFERENCE			
AUTHORS Contact: Rod wing			
Arizona Genomics Institute			
University of Arizona			
Biological Sciences West, 448A, P.O. Box 210088, Tucson, AZ			
85721-0088, USA			
Tel: 520 626 3967			
Fax: 520 621 9288			
Email: http://genome.arizona.edu			
PCR Primers			
FORWARD: gta aaa cga cgg cca gtg			
BACKWARD: gga aac agc tat gac cat g			
Plate: 15 row: 0 column: 18			
Seq primer: gga aac agc tat gac cat g.			
Location/Qualifiers			
1..794			
/organism="Oryza sativa (japonica cultivar-group)"			
/mol_type="mRNA"			
/cultivar="Nipponbare"			
/db_xref="taxon:39947"			
/clone="OSJNEB15018"			
/tissue_type="Leaf"			
/dev_stage="3 week"			
/lab_host="DH10B"			
/clone_lib="OSJNEB"			
/note="Vector: pBluescript II KS +; Site_1: EcoRI; Site_2:			
XhoI; 24 hrs after inoculation with Rice Blast (Che			
86061)"			
ORIGIN			
Query Match 9.5%; Score 65.2; DB 14; Length 794;			
Best Local Similarity 49.7%; Pred. No. 6.9;			
Matches 197; Conservative 0; Mismatches 193; Indels 6; Gaps 1;			
Oy	74	TCAAGTGAGGGCGAGGGCGACGGCAACCCCTTGGCCGGACCCAGAGCATGAATCC	133
Db	793	TCATGACGTCGTCGACGTCGCCGACGACGAGCGCGCTGCCGATCAAGAGCAGCAGTACC	734
Oy	134	ACGTGACCGAGGGCGCCCCCTGCCCTTGCCCTTGACATCTGGCCCCCTGCTCGAGT	193
Db	733	GCAGCGCGAGCTTCTACACCATGGGGCGCGGCTACGTCAACCGTCCCGCGCTCGACC	674
Oy	194	ACGGCAGCAGGACCTTCGTGCAACCAACCGCCGAGATCCCAGATTCTTCAAGCAGACT	253
Db	673	CGGGCCTCGTCTACGACTCCACACCAAGACTACATGCTTAACCTTGCGGCTCGGCA	614
Oy	254	TCCCGAGGGCTTCACTGGGAGAAGAACCAACCACTACGAGGACGGCGCATCTGACCG	313
Db	613	TGGCGACGACGGCGGTGAAGAGATTAACCAACCGCGCTCTCTGCGCCAAAGCTCAAG	554

[illegible]

```

RESULT 11
BU626888/c 697 bp mRNA linear EST 23-SEP-2002
LOCUS
DEFINITION UI-H-FT0-bhn-d-01-0-UI.s1 NCI CGAP_FTO Homo sapiens cDNA clone
UI-H-FT0-bhn-d-01-0-UI 3', mRNA sequence.
ACCESSION BU626888
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 697)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Robeff-Pamela, U of I
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES
source
location/Qualifiers
1..697
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FT0-bhn-d-01-0-UI"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI_CGAP_FTO"
/notes="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: Ecor I; Site 2: Not I;
NCI CGAP_FTO is a cDNA library constructed from a pool of
81 RNA samples from Alveolar Macrophages challenged with
different treatments. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GGCCATGCCG. The cell line
was provided by Gary W. Hunninghake from the University of
Iowa.
TAG_TISSUE=Human Lung Alveolar Macrophage
TAG_LIB=UI-H-FT0
TAG_SEQ=GGCCATGCCG"

ORIGIN
Query Match 9.4%; Score 64.8; DB 13; Length 697;
Best Local Similarity 52.6%; Pred. NO. 7.6;

```

	Matches	141;	Conservative	0;	Mismatches	127;	Indels	0;	Gaps	0;
QY	214	CACCACACCGCGAGATCC	CCCCCACTTCTTCAAGCAGAGCTT	CCCCCGAGGCGTTCA	CACTGG	273				
DB	671	CCCGACACATGAAGCAGC	ACGACTTCTTCAAGTCCGCCATG	CCCGAAGGTAA	CGTCCAG	612				
QY	274	GAGAGAAACCACTTA	CGAGGACGGCGGCATCTGA	CCGCCCA	CGAGACCAAGCCTG	333				
DB	611	GAGCGCACCATCTTCTT	CAAGGACCA	CGGCACTTCAAGAC	CCCGCGCGAGGTGAAGTTG	552				
QY	334	GAGGGCAACTGCTGAT	CTACAAGGTGAAGGTGCA	CGGCACCACTT	CCCCCGCAGCGGC	393				
DB	551	GAGGGCGACACCTCTG	TGAACCCGACTCGAGCTGA	AGGGGCATCGACTT	CAAGGAGGACGGC	492				
QY	394	CCCGTGATGAAGACA	GAGCGCGGCTGGAGGCCAGCACCG	AGGTGTGTA	CCCCGAG	453				
DB	491	AACATCTCTGGGGCACA	GCTGGAGTACAACTACA	CAGCCACA	AGTCTATATCA	TATGACC	432			
QY	454	AACGGCGTGCTGTG	CGCGCCGGAACGTGA	481						
DB	431	GACAAGCAGAGAAGAC	GCATC	AAAGGTGA	404					

RESULT	12
CK199961	
LOCUS	821 bp mRNA linear EST 08-DEC-2003
DEFINITION	FGAS008468 Triticum aestivum FGAS: Library 3 Gate 6 Triticum aestivum cDNA, mRNA sequence.
ACCESSION	CK199961
VERSION	CK199961.1 GI:39562351

ORGANISM	Triticum aestivum	(bread wheat)
SOURCE	Triticum aestivum (bread wheat)	
RECORDS	Triticum aestivum (bread wheat)	

Eubalyptus; Vitaceae; Eucaryota; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.

REFERENCE  
AUTHORS  
1 (bases 1 to 821)  
Allard, F., Crosby, W. L., Danyluk, J., Eudes, F., Frick, M., Gaudet, D., ...

**TITLE** Functional Genomics of Abiotic Stresses In Wheat and Canola Crops  
**JOURNAL** Unpublished (2003)  
**COMMENT** Contact: Wm L Crosby

BIOINFORMATICS  
 University of Saskatchewan, Department of Computer Science  
 1C101 Engineering Building, 57 Campus Drive, Saskatoon,  
 Saskatchewan, S7N 5A9, Canada  
 Tel: 306 966 1769  
 Fax: 306 966 2033  
 Email: fgas\_estecs.usask.ca  
 This sequence is the direct result of the Base calling software  
 Phred (default parameters). It is the raw base calls. To aid in the  
 identification of the high quality insert the software Lucy  
 (default parameters) has been run on this sequence. Lucy identified  
 the region [101,701].  
 Plate: L3C111 row: K column: 01.

## FEATURES

**Source**

1. 821  
/organism="Triticum aestivum"  
/mol\_type="mRNA"  
/db\_xref="taxon:4565"  
/clone\_lib="Triticum aestivum FGAS: Library 3 Gate 6"  
/note="Organ: Root; Vector: pCMV.SPORT6; Root tissue from  
control, cold-acclimated and salt stressed wheat cultivar  
Norstar. 7 mRNA populations were combined before  
constructing the library; 7 day non-acclimated roots, 1,  
23, and 53 days cold-acclimated at 4C, and 30 minutes, 3  
hours and 6 hours treated roots with 200mM NaCl.  
Non-acclimated and cold-acclimated plants were grown in  
vermiculite while salt stressed plant were grown  
hydropotonically. First strand synthesis in this library was  
done in the presence of methylated dCTP thereby protecting

from internal cleavage with NotI."

Query Match	9.4%;	Score 64.6;	DB 14;	Length 821;
Best Local Similarity	48.1%;	Pred. No. 8.4;		
Matches 215;	Conservative	0;	Mismatches 229;	Indels 3;
				Gaps 1,

[illegible]

RESULT 13	BI386976	LOCUS	BI386976	613 bp	mRNA	linear	EST 26-AUG-2003
DEFINITION	BFL26 001679	Amphioxus 26hr CDNA library (Name convention: BFL26 CDNA library)	Amphioxus 26hr CDNA library (Name convention: BFL26 CDNA library)	613 bp	mRNA	linear	EST 26-AUG-2003

ACCESSION	BI386976	GI:30921865
VERSION	BI386976.1	
KEYWORDS	EST.	
SOURCE	Branchiostoma floridae (Florida lancelet)	
ORGANISM	Branchiostoma floridae	

REFERENCE  
AUTHORS

1 (bases 1 to 613)  
panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J.,

**TITLE** Herwig, R., Vingron, M. and Lehrach, H.  
New evidence for genome-wide duplications at the origin of

vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL	Genome Res. 13 (6A), 1056-1066 (2003)
MEDLINE	22683279
PUBMED	12799346
COMMENT	Contact: Panopoulou G

laboratory 145, dept. Lehrach  
Max-Planck-Institut fuer Molekulare Genetik  
Inhestr. 63-73, D-14195 Berlin, Germany  
Tel: +49 30 8413 1235  
Fax: +49 30 8413 1128  
Email: panopoulos@molgen.mpg.de  
The library was characterised by oligonucleotide fingerprinting  
(ONFP) to reduce sequencing redundancy. According to the ONFP  
procedure, clones giving the same hybridisation pattern with a  
battery of 200 8mer oligonucleotides are grouped into clusters. One

clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/> Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCAGGCTTACACTTATGCTCCGGCTCG 3' (M13RSP)

BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGGTCCGGAATCCCGGGT-3' pSport3/86

High quality sequence stop: 613.

FEATURES

source

1. .613

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMGp531N0799"

/tissue\_type="whole embryo"

/dev\_stage="26 hrs (neurula stage)"

/lab\_host="E.coli, XLI blue"

/clone\_lib="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531)"

/note="Vector: pSport1 (Gibco BRL); Site\_1: Sali, KpnI, EcoRI (5'); Site\_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGCGCCC (T)15-3' and a Sali 5'- TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."

ORIGIN

Query Match

Best Local Similarity 47.6%; Pred. No.9.1;

Matches 189; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

9.3%; Score 64.2; DB 12; Length 613;

QY 36 CAAGATGTACATGAGGGCACCGTGAACGGCCACTACTTCAAGTGCAGGGCGAGGGCGA 95

Db 88 CGACATCCACCTTCACGGCTCCATCAACGGCCACGAGTTCGACATGTGGGGGAGGAAA 147

QY 96 CGGCACCCCTTCGCCGGCACCGAGCATGAAATCCAGTGAACCGAGGGCGCCCCCT 155

Db 148 AGCGGACCCGGAACGCGGCTCGTGGTGACCACAGCGAAATCCACCAAGGGTGCCCTGAA 207

QY 156 GCCCTTCGCTTGACATCCTGGCCCCCTGCTGCGAGTACGGCAGACACTTCGTGCA 215

Db 208 GTTCTCTCCCTACTGATGATCCCCACCTCGGGTACGGGTACTACCACTACCTCCCCA 267

QY 216 CCACACCGCCGAGATCCCGCACTTCTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGA 275

Db 268 CCCGAGCGGACCCCTCGCCTTCCAGGCCCTCCATGTGGAAGGATCGGGGTATGCAGTCTA 327

QY 276 GAGAACCAACCACTACGAGAGCGCGGCATCCTGACCGCCCAACGAGACACCAAGCCTTGA 335

Db 328 CCGCGTGTTCGACTTTGAAGACGAGGCAAGCTGACTACCGAGTTTAAGTACTCTTACGA 387

QY 336 GGGCAACTGCCTGATCTACAAGGTGAAGGTGCACGGCACCACTTCCCGCGAGGGGCC 395

Db 388 GGGTTCCTCATATCAAGGCCGACATGAAGCTGATGGAAAGCGTTTCCCTGATGACGGCC 447

QY 396 CGTGATGAAGAACAAGAGCGCGCTGGAGGCCAGC 432

Db 448 AGTCATGACCAGCCAGATTTGTGACCAAGACGCGCTGC 484

RESULT 14

BI386699 643 bp mRNA linear EST 26-AUG-2003

LOCUS BFL26\_001402 Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531) Branchiostoma floridae cDNA clone MPMGp531L1370 5', mRNA sequence.

ACCESSION BI386699

VERSION BI386699.1 GI:30921604

KEYWORDS

SOURCE EST.

ORGANISM Branchiostoma floridae (Florida lancelet)

REFERENCE Branchiostoma floridae

AUTHORS Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae; Branchiostoma.

TITLE 1 (bases 1 to 643)

JOURNAL Panopoulou,G., Hennig,S., Groth,D., Krause,A., Poustka,A.J., Herwig,R., Vingron,M. and Lehrach,H.

MEDLINE New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

PUBMED Genome Res. 13 (6A), 1056-1066 (2003)

COMMENT 12799346

Contact: Panopoulou G

Laboratory 145, dept.Lehrach

Max-Planck-Institut fuer Molekulare Genetik

Innestr.63-73, D-14195 Berlin, Germany

Tel: +49 30 8413 1235

Fax: +49 30 8413 1128

Email: panopoul@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 8mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/> Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)

PCR Primers

FORWARD: 5' CCCAGGCTTACACTTATGCTTCCGGCTCG 3' (M13RSP)

BACKWARD: 5' GCTATTACGCCAGCTGGCGAAGGGGATGTG 3' (M13FSP)

Insert Length: 1 Std Error: 0.00

Seq primer: 5'-CCGGTCCGGAATTCCTCCGGGT-3' pSport3/86

High quality sequence stop: 643.

FEATURES

source

1. .643

/organism="Branchiostoma floridae"

/mol\_type="mRNA"

/db\_xref="taxon:7739"

/clone="MPMGp531L1370"

/tissue\_type="whole embryo"

/dev\_stage="26 hrs (neurula stage)"

/lab\_host="E.coli, XLI blue"

/clone\_lib="Amphioxus 26hr cDNA library (Name convention: BFL26 or MPMGp531)"

/note="Vector: pSport1 (Gibco BRL); Site\_1: Sali, KpnI, EcoRI (5'); Site\_2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pGACTAGTTCTAGATCGGAGCGCGCCC (T)15-3' and a Sali 5'- TCGACCCACGCGTCCG-3'adapters (Gibco BRL)."

ORIGIN

Query Match

Best Local Similarity 47.6%; Pred. No.9.2;

Matches 189; Conservative 0; Mismatches 208; Indels 0; Gaps 0;

9.3%; Score 64.2; DB 12; Length 643;

QY 36 CAAGATGTACATGAGGGCACCGTGAACGGCCACTACTTCAAGTGCAGGGCGAGGGCGA 95

Db 77 CGACATCCACCTTCACGGCTCCATCAACGGCCACGAGTTTCGACATGTGGGGGAGGAAA 136

QY 96 CGGCACCCCTTCGCCGGCACCCAGAGCATGAAATCCAGTGACCGAGGGCGCCCCCT 155

Db 137 AGCGGACCCGGAACGCGGCTCGGTGTGACCAACGGAATCCACCAAGGGTCCCTGAA 196

QY 156 GCCCTTCGCTTGACATCCTGGCCCCCTGCTGCGAGTACGGCAGCAGACCTTCGTGCA 215

Db 197 GTTCTCTCCCTACTGATGATCCCCACCTCGGGTACGGGTACTACCACTACCTCCCCCTA 256





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 15, 2004, 10:05:08 ; Search time 395 Seconds  
(without alignments)  
7388.638 Million cell updates/sec

Title: US-09-976-673-11

Perfect score: 687

Sequence: 1 atggtgagcgccctgctgaa.....tgcccgagaagcgccaactga 687

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	687	100.0	687	6	ABL41172 H. crispa
2	671	97.7	687	6	ABL41173 H. crispa
3	671	97.7	1396	6	ABL41174 Cr-449-ta
4	671	97.7	1396	7	ACA61024 DNA encod
5	669.4	97.4	1424	6	ABL41175
6	669.4	97.4	1424	7	ACA61025 DNA encod
7	446.6	65.0	1376	7	ACA61026
8	441.2	64.2	681	6	ABL41170 H. crispa
9	439.6	64.0	681	6	ABL41171 H. crispa
10	438.2	63.8	684	6	ABL41180 H. crispa
11	436.6	63.6	684	6	ABL41169 H. crispa
12	436.6	63.6	760	7	ABA00802
13	436.6	63.6	760	7	ABA00806
14	435	63.3	910	6	ABL41167 H. crispa
15	435	63.3	910	6	ABL41182 H. crispa
16	431.8	62.9	760	7	ABA00805
17	431.8	62.9	760	7	ABA00804
18	431.8	62.9	760	7	ABA00803
19	431.8	62.9	908	6	ABL41168 H. crispa
20	429.2	62.5	680	6	ABL41181 H. crispa
21	422.6	61.5	696	3	AAA50885 A. sulcat
22	417.8	60.8	707	6	AAD46286
23	417.8	60.8	1398	7	ACA61028 DNA encod

24	415.6	60.5	1404	7	ACA61027	ACA61027	DNA encod
25	415.2	60.4	699	6	AAD46287	Aad46287	Anemonia
26	362.6	52.8	681	6	ABA00245	Aba00245	C. gigant
27	361.2	52.6	835	6	ABA00244	Aba00244	C. gigant
28	306	44.5	678	9	ADC24132	Adc24132	Discosoma
29	298.8	43.5	681	4	AAH47656	Aah47656	Anthozoan
30	298.2	43.4	681	9	ADC24130	Adc24130	Discosoma
31	297.4	43.3	723	7	ABZ22476	Abz22476	Mammalian
32	292.4	42.6	678	6	AAD46278	Aad46278	Discosoma
33	292.4	42.6	678	6	AAD28208	Aad28208	Discosoma
34	292.4	42.6	678	6	AAD28207	Aad28207	Discosoma
35	292.4	42.6	695	3	AAA48743	Aaa48743	Humanised
36	291.8	42.5	678	9	ADC24128	Adc24128	Discosoma
37	291.4	42.4	681	9	ADC24134	Adc24134	Discosoma
38	291.4	42.4	4692	6	AAL47954	Aal47954	Modified
39	291.4	42.4	4692	7	ACC44640	Acc44640	Vector pD
40	291.4	42.4	7910	4	AAD09979	Aad09979	pBit(dHSP
41	291.4	42.4	9320	6	ABS56664	Abs56664	Plasmid p
42	290.8	42.3	6893	9	ADE24111	Ade24111	Proviral
43	289.8	42.2	5436	4	AAD10003	Aad10003	Plasmid p
44	289.2	42.1	678	6	AAD46282	Aad46282	Discosoma
45	289.2	42.1	678	6	AAD28209	Aad28209	Discosoma

ALIGNMENTS

RESULT 1	
ABL41172	
ID	ABL41172 standard; cDNA; 687 BP.
XX	
AC	ABL41172;
XX	
DT	12-AUG-2002 (first entry)
XX	
DE	H. crispa fluorescent protein mutant FP10-crl encoding cDNA.
XX	
KW	Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;
KW	colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;
KW	fluorescence resonance energy transfer; gene expression; mutant; ss.
XX	
OS	Heteractis crispa.
XX	
FH	Key
FT	CDS
FT	1..687
FT	/*tag= a
FT	/product= "mutant fluorescent protein"
XX	
PN	WO200230965-A2.
XX	
PD	18-APR-2002.
XX	
PF	12-OCT-2001; 2001WO-US032080.
XX	
PR	12-OCT-2000; 2000US-0240018P.
PR	16-JUL-2001; 2001US-0306131P.
XX	
PA	(CLON-) CLONTECH LAB INC.
XX	
PI	Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;
XX	
DR	WPI; 2002-444170/47.
DR	P-PSDB; ABB07994.
XX	
PT	Novel nucleic acid encoding Stichodactylidae chromoprotein and its
PT	fluorescent mutant useful as coloring agent, labels in analyte detection
PT	assays, markers in recombinant DNA applications and filters in
PT	sunscreens.
XX	
PS	Claim 5; Fig 10; 81pp; English.
XX	
CC	The invention relates to a nucleic acid present in other than its natural
CC	environment and encoding an Stichodactylidae chromoprotein or its

CC fluorescent mutant, where the fluorescent protein has an emission maximum  
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are  
CC useful in applications employing a chromo or fluorescent nucleic acid or  
CC protein. Recombinant vectors comprising the nucleic acid is useful for  
CC producing an Anthozoan chromo and/or fluorescent protein. The  
CC chromoproteins, and their fluorescent mutants are useful as colouring  
CC agents capable of imparting colour or pigment to a particular composition  
CC of matter. The chromoproteins can be incorporated into a variety of  
CC different compositions including food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g. animals and plants, and as labels in  
CC analyte detection assays, e.g. assays for biological analytes of interest  
CC (see ABL41167 for a detailed description of the various uses of the  
CC chromoproteins). The present sequence represents the H. crispa  
CC fluorescent protein mutant FP10-crl encoding cDNA

XX  
SQ Sequence 687 BP; 150 A; 232 C; 212 G; 93 T; 0 U; 0 Other;

Query Match 100.0%; Score 687; DB 6; Length 687;  
Best Local Similarity 100.0%; Pred. No. 1e-111;  
Matches 687; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGTGAGCGGCGCTGTGAAGAGAGATGCGCATCAAGATGTACATGAGGGCACCGTG 60  
Db 1 ATGGTGAGCGGCGCTGTGAAGAGAGATGCGCATCAAGATGTACATGAGGGCACCGTG 60  
QY 61 AACGGCCACTACTTCAAGTGCAGAGGGCGGCGGCAACCCCTTCGCCGGCACCCAG 120  
Db 61 AACGGCCACTACTTCAAGTGCAGAGGGCGGCGGCGGCAACCCCTTCGCCGGCACCCAG 120  
QY 121 AGCATGAGAATCCAGGTGACCGAGGGCGGCCCTGACCTTGACCTTGACATCCTGGCC 180  
Db 121 AGCATGAGAATCCAGGTGACCGAGGGCGGCCCTGACCTTGACCTTGACATCCTGGCC 180  
QY 181 CCCTGCTGCAGTAGCGGAGCAGGACCTTGTGACACCAACCGCGAGATCCCGACTTC 240  
Db 181 CCCTGCTGCAGTAGCGGAGCAGGACCTTGTGACACCAACCGCGAGATCCCGACTTC 240  
QY 241 TTCAAGCAGAGCTTCCCGGAGGGCTTCACTCTGGGAGAGAACCAACCACTACGAGACGGC 300  
Db 241 TTCAAGCAGAGCTTCCCGGAGGGCTTCACTCTGGGAGAGAACCAACCACTACGAGACGGC 300  
QY 301 GGCATCCTGACCGCCACCAAGACACCAAGCTTGAGGGGCAACTGCTGATCTACAAGGTG 360  
Db 301 GGCATCCTGACCGCCACCAAGACACCAAGCTTGAGGGGCAACTGCTGATCTACAAGGTG 360  
QY 361 AAGGTGACCGGCACCAACTTCCCCGCGGAGGGCCCCGTGATGAAGAACAAGACGGCGGC 420  
Db 361 AAGGTGACCGGCACCAACTTCCCCGCGGAGGGCCCCGTGATGAAGAACAAGACGGCGGC 420  
QY 421 TGGGAGCCCAAGCAGCGAGGTGTATCCCCGAGAACGGCGTGTGCGGCCGGAACGTG 480  
Db 421 TGGGAGCCCAAGCAGCGAGGTGTATCCCCGAGAACGGCGTGTGCGGCCGGAACGTG 480  
QY 481 ATGGCCCTGAAGGTGGGCGACCGGACCTGATCTGCCACCACTACACCACTACCGGAGC 540  
Db 481 ATGGCCCTGAAGGTGGGCGACCGGACCTGATCTGCCACCACTACACCACTACCGGAGC 540  
QY 541 AAGAGGCGCTGCGCGCCCTGACCATGCGCGGCTTCCACTTACCGACATCCGGCTCCAG 600  
Db 541 AAGAGGCGCTGCGCGCCCTGACCATGCGCGGCTTCCACTTACCGACATCCGGCTCCAG 600  
QY 601 ATGCTGCGGAGAGAGAGAGAGAGTACTTTCGAGCTGTACGAGGCCAGCGGTGGCTAC 660  
Db 601 ATGCTGCGGAGAGAGAGAGAGAGTACTTTCGAGCTGTACGAGGCCAGCGGTGGCTAC 660  
QY 661 AGCGACCTGCCCCGAGAAGGCCAACTGA 687  
Db 661 AGCGACCTGCCCCGAGAAGGCCAACTGA 687

RESULT 2  
ABL41173  
ID ABL41173 standard; cDNA; 687 BP.

XX  
AC ABL41173;  
XX  
XX 12-AUG-2002 (first entry)  
DT  
XX  
DE H. crispa alternative fluorescent protein mutant FP10-crl cDNA.  
XX  
KW Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;  
KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;  
KW fluorescence resonance energy transfer; gene expression; mutant; ss.  
XX  
OS Heteractis crispa.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..687  
FT misc\_feature 4..6  
FT /tag= a  
FT /product= "mutant fluorescent protein"  
FT /tag= b  
FT /note= "the amino acid Val encoded by the above codon is  
FT not indicated in the corresponding protein"

XX WO200230965-A2.  
XX  
XX 18-APR-2002.  
XX  
XX 12-OCT-2001; 2001WO-US032080.  
XX  
XX 12-OCT-2000; 2000US-024001BP.  
XX  
XX 16-JUL-2001; 2001US-0306131P.  
XX  
XX (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;  
XX  
XX WPI; 2002-444170/47.  
DR P-PSDB; ABB07995.  
DR  
XX  
XX Novel nucleic acid encoding Stichodactylidae chromoprotein and its  
PT fluorescent mutant useful as coloring agent, labels in analyte detection  
PT assays, markers in recombinant DNA applications and filters in  
PT sunscreens.  
PT  
XX  
XX Claim 5; Fig 10; 81bp; English.

CC The invention relates to a nucleic acid present in other than its natural  
CC environment and encoding an Stichodactylidae chromoprotein or its  
CC fluorescent mutant, where the fluorescent protein has an emission maximum  
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are  
CC useful in applications employing a chromo or fluorescent nucleic acid or  
CC protein. Recombinant vectors comprising the nucleic acid is useful for  
CC producing an Anthozoan chromo and/or fluorescent protein. The  
CC chromoproteins, and their fluorescent mutants are useful as colouring  
CC agents capable of imparting colour or pigment to a particular composition  
CC of matter. The chromoproteins can be incorporated into a variety of  
CC different compositions including food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g. animals and plants, and as labels in  
CC analyte detection assays, e.g. assays for biological analytes of interest  
CC (see ABL41167 for a detailed description of the various uses of the  
CC chromoproteins). The present sequence represents the H. crispa  
CC alternative fluorescent protein mutant FP10-crl encoding cDNA

XX  
SQ Sequence 687 BP; 145 A; 230 C; 219 G; 93 T; 0 U; 0 Other;

Query Match 97.7%; Score 671; DB 6; Length 687;  
Best Local Similarity 98.5%; Pred. No. 6.5e-109;  
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGTGAGCGGCGCTGTGAAGAGAGAGTATGCGCATCAAGATGTACATGAGGGCACCGTG 60  
Db 1 ATGGTGAGCGGCGCTGTGAAGAGAGAGATGCGCATCAAGATGTACATGAGGGCACCGTG 60  
QY 61 AACGGCCACTACTTCAAGTGCAGAGGGCGGCGGCAACCCCTTCGCCGGCACCCAG 120

```

Db      61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGACGGCAACCTCTTCCGGCACCCAG 120
Qy      121 AGCATGAGAATCCACGTGACCGAGGGCGCCCCCTGCGCTTGCCTTGACATCTGGCG 180
Db      121 AGCATGGCGATCCACGTGACCGAGGGGGCCCCCTGCGCTTGCCTTGACATCTGGCG 180
Qy      181 CCCTGCTGCCAGTACGGCAGCAGACCTTCTGTGACCAACACCGCCGAGATCCCGCACTTC 240
Db      181 CCCTGCTGCCAGTACGGCAGCAGACCTTCTGTGACCAACACCGCCGAGATCCCGCACTTC 240
Qy      241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGAGAGAACCACTTACGAGGACGGC 300
Db      241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGAGAGAACCACTTACGAGGACGGC 300
Qy      301 GGCATCTGACCGCCCAACGAGCAGCCTGTGAGGGCAACTGCTGATCTACAAGGTG 360
Db      301 GGCATCTGACCGCCCAACGAGCAGCCTGTGAGGGCAACTGCTGATCTACAAGGTG 360
Qy      361 AAGGTGCACGGCACCACTTCCCGCGCAGCGGCCCGTGTATGAGAACAAGAGCGCGGC 420
Db      361 AAGGTGTGGGCACCACTTCCCGCGCAGCGGCCCGTGTATGAGAACAAGAGCGCGGC 420
Qy      421 TGGAGCCCAACGACCGAGGTGTATCCCCGAGAACGGCGTGTGCGCGCGGAACGTG 480
Db      421 TGGAGCCCAACGACCGAGGTGTATCCCCGAGAACGGCGTGTGCGCGCGGAACGTG 480
Qy      481 ATGGCCCTGAAGGTGGGGCAGCCGGCACTGTATGCGCAACCACTACCAAGCTACCGGAGC 540
Db      481 ATGGCCCTGAAGGTGGGGCAGCCGGCGCTGATCTGCCAACCACTACCAAGCTACCGGAGC 540
Qy      541 AAGAAAGCCGTGCGCGCCCTGACCATGCGGGCTTCCACTTCAACGACATCCGGCTCCAG 600
Db      541 AAGAAAGCCGTGCGCGCCCTGACCATGCGGGCTTCCACTTCAACGACATCCGGCTCCAG 600
Qy      601 ATGCTCGGAAGAAGAAGAGAGACTTTCGAGCTGTACGAGGCCAGCGTGCGCCGGTAC 660
Db      601 ATGCTCGGAAGAAGAAGAGAGACTTTCGAGCTGTACGAGGCCAGCGTGCGCCGGTAC 660
Qy      661 AGCGACTGCCCCGAGAAGGCCAATTGA 687
Db      661 AGCGACTGCCCCGAGAAGGCCAATTGA 687

```

```

RESULT 3
ABL41174
ID      ABL41174 standard, DNA; 1396 BP.
XX
AC      ABL41174;
XX
DT      12-AUG-2002 (first entry)
XX
DE      Cr-449-tandem fusion protein nucleotide sequence.
XX
KW      Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;
KW      colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;
KW      fluorescence resonance energy transfer; fusion protein; ds.
XX
OS      Synthetic.
OS      Heteractis crispa.
XX
FH      Key
FT      CDS
FT      14..1396
FT      /*tag= a
FT      /product= "Cr-449-tandem fusion protein"
FT      misc_feature
FT      698..709
FT      /*tag= b
FT      /note= "linker nucleotide fragment"
XX
PN      WO200230965-A2.
XX
PD      18-APR-2002.
XX

```

```

PF      12-OCT-2001; 2001WO-US032080.
XX
PR      12-OCT-2000; 2000US-0240018P.
PR      16-JUL-2001; 2001US-0306131P.
XX
PA      (CLON-) CLONTECH LAB INC.
XX
PI      Lukyanov SA, Fradkov AF, Lukyanov KA, Gurekaya NG;
XX
DR      WPI; 2002-444170/47.
DR      P-PSDB; ABB07996.
XX
PT      Novel nucleic acid encoding Stichodactylidae chromoprotein and its
PT      fluorescent mutant useful as coloring agent, labels in analyte detection
PT      assays, markers in recombinant DNA applications and filters in
PT      sunscreens.
PS      Disclosure; Fig 12; 81pp; English.
XX
CC      The invention relates to a nucleic acid present in other than its natural
CC      environment and encoding an Stichodactylidae chromoprotein or its
CC      fluorescent mutant, where the fluorescent protein has an emission maximum
CC      ranging from 580-660 nm. The polynucleotides and encoded proteins are
CC      useful in applications employing a chromo or fluorescent nucleic acid or
CC      protein. Recombinant vectors comprising the nucleic acid is useful for
CC      producing an Anthozoan chromo and/or fluorescent protein. The
CC      chromoproteins, and their fluorescent mutants are useful as colouring
CC      agents capable of imparting colour or pigment to a particular composition
CC      of matter. The chromoproteins can be incorporated into a variety of
CC      different compositions including food compositions, pharmaceuticals,
CC      cosmetics, living organisms, e.g. animals and plants, and as labels in
CC      analyte detection assays, e.g. assays for biological analytes of interest
CC      (see ABL41167 for a detailed description of the various uses of the
CC      chromoproteins). The present sequence represents the H. crispa Cr-449-
CC      tandem fusion protein nucleotide sequence
XX
SQ      Sequence 1396 BP; 293 A; 471 C; 444 G; 188 T; 0 U; 0 Other;

```

```

Query Match      97.7%; Score 671; DB 6; Length 1396;
Best Local Similarity 98.5%; Pred. No. 6.5e-109;
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```

```

Qy      1 ATGCTAGCGGCGCTGTGAAGAGAGATGCGCATCAAGTGTACATGAGGGCACTGTG 60
Db      710 ATGCTAGCGGCGCTGTGAAGAGAGATGCGCATCAAGTGTACATGAGGGCACTGTG 769
Qy      61 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGAGCGCAACCCCTTGCSCGGACCCAG 120
Db      770 AACGGCCACTACTTCAAGTGCAGGGCGAGGGCGAGCGCAACCCCTTGCSCGGACCCAG 829
Qy      121 AGCATGAGATCCACGTGACCGAGGGGGCCCCCTTGCCTTGCACATCTTGGCC 180
Db      830 AGCATGCGGATCCACGTGACCGAGGGGGCCCCCTTGCCTTGCACATCTTGGCC 889
Qy      181 CCCTGCTGCGAGTACGGCAGGAGGACCTTGTGACCAACCGCGAGATCCCGCACTTC 240
Db      890 CCCTGCTGCGAGTACGGCAGGAGGACCTTGTGACCAACCGCGAGATCCCGCACTTC 949
Qy      241 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGAGAGAACCACTTACGAGGACGGC 300
Db      950 TTCAAGCAGAGCTTCCCGAGGGCTTCACTTGGAGAGAACCACTTACGAGGACGGC 1009
Qy      301 GGCATCTGACCGCCCAACGAGCAGCAGCCTTGAAGGCAACTGCTGATCTACAAGGTG 360
Db      1010 GGCATCTGACCGCCCAACGAGCAGCAGCCTTGAAGGCAACTGCTGATCTACAAGGTG 1069
Qy      361 AAGGTGCACGGCACCACTTCCCGCGCAGCGGCCCGTGTATGAAGAACAGAGCGGGCGC 420
Db      1070 AAGGTGCTGGGCACCACTTCCCGCGCAGCGGCCCGTGTATGAAGAACAGAGCGGGCGC 1129
Qy      421 TGGGAGCCCAAGCAGCGAGGTGTATCCCGAGAACGGCGTGTGTGCGGCGCGAACGTG 480
Db      1130 TGGGAGCCCAAGCAGCGAGGTGTATCCCGAGAACGGCGTGTGTGCGGCGCGAACGTG 1189

```

QY 481 ATGCCCTGAAGGTGGGCGACCGGACCTGATCTGGCCACCCTACACAGCTACCGGAGC 540  
CC |||||||  
XX 1190 ATGGCCTGAAGGTGGGCGACCGGCGGTGATCTGGCCACTACACAGCTACCGGAGC 1249  
QY 541 AAGAGGCCGTGCGCGCCTGACCATGCGCGCTTCCACTTCAACCGACATCCGGCTCCAG 600  
CC |||||||  
Db 1250 AAGAGGCCGTGCGCGCCTGACCATGCGCGCTTCCACTTCAACCGACATCCGGCTCCAG 1309  
QY 601 ATGCTCGGGAAGAAGAGGACGAGTACTTGGAGCTGTACGAGGCCAGCGTGCGCCGGTAC 660  
CC |||||||  
Db 1310 ATGCTCGGGAAGAAGAGGACGAGTACTTGGAGCTGTACGAGGCCAGCGTGCGCCGGTAC 1369  
QY 661 AGCGACCTGCGCGGAGAGGCCCACTGA 687  
CC |||||||  
Db 1370 AGCGACCTGCGCGGAGAGGCCCACTGA 1396

RESULT 4  
ACA61024  
ID ACA61024 standard; DNA; 1396 BP.  
XX  
AC ACA61024;  
XX  
DT 09-JUL-2003 (first entry)  
XX  
DE DNA encoding chromoprotein Cr-449-tandem.  
XX  
KM Chromo/fluorescent domain; labeled fusion protein;  
KW site-specific gene modification; chromoprotein; colouring agent;  
XX food composition; pharmaceutical; cosmetic; Cr-449-tandem; gene; ds.  
OS Anthozoa.

Key Location/Qualifiers  
FT CDS 14..1396  
FT /\*tag= a  
FT /product= "Cr-449-tandem"

PN WO2003031590-A2.  
XX  
PD 17-APR-2003.  
XX  
PE 10-OCT-2002; 2002WO-US032560.  
XX  
PR 12-OCT-2001; 2001US-00976673.  
PR 11-FEB-2002; 2002US-0356225P.  
PR 22-MAY-2002; 2002US-0383336P.

PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA;  
XX  
DR WPI; 2003-381709/36.  
DR P-PSDB; ABU09921.

PT New nucleic acid encoding polypeptide products having at least two linked  
PT chromo/fluorescent domains; useful for generating transgenic plants or  
PT animals or site-specific gene modifications in cell lines.  
XX

PS Disclosure; Fig 1; 68bp; English.

XX The invention describes a nucleic acid encoding a polypeptide product  
CC comprising a first and a second chromo/fluorescent domain, optionally  
CC joined by a linking domain. The first and second chromo/fluorescent  
CC domains associate with each other under intracellular conditions so that  
CC the encoded polypeptide assumes a tertiary structure. The nucleic acid  
CC and the protein are useful in producing labeled fusion proteins that have  
CC a precise and predictable signal to fusion partner ratio. The nucleic  
CC acid may also be used in generating transgenic, non-human plants or  
CC animals or site-specific gene modifications in cell lines. The  
CC chromoproteins may be used as colouring agents, as a food composition, in  
CC pharmaceuticals or cosmetics, as labels in analyte detection assays or as

CC selectable markers in recombinant DNA applications. This sequence encodes  
CC chromo/fluorescent domain fusion protein Cr-449-tandem  
XX  
SQ Sequence 1396 BP; 293 A; 471 C; 444 G; 188 T; 0 U; 0 Other;

Query Match 97.7%; Score 671; DB 7; Length 1396;  
Best Local Similarity 98.5%; Pred. No. 6.5e-109;  
Matches 677; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 ATGGTAGCGGCGCTGTGAAGAGAGATATGCGCATCAAGATGTACATGAGGCAACCGTG 60  
CC |||||||  
Db 710 ATGGTAGCGGCGCTGTGAAGAGAGATATGCGCATCAAGATGTACATGAGGCAACCGTG 769  
QY 61 AAGGCCACTTACTTCAAGTGCAGAGGCGGCGGCGACCGCAACCCCTTGGCGGCACCCAG 120  
CC |||||||  
Db 770 AAGGCCACTTACTTCAAGTGCAGAGGCGGCGGCGACCGCAACCCCTTGGCGGCACCCAG 829  
QY 121 AGCATGAAGATCCACGTTGACCGGAGGCGCGCCCTTGCCCTTGCAATCTGCGCC 180  
CC |||||||  
Db 830 AGCATGCGGATCCACGTTGACCGGAGGCGCGCCCTTGCCCTTGCAATCTGCGCC 889  
QY 181 CCCTGCTGGAGTACCGGACGACGACGACCTTGTGCAACCAACCGCGAGATCCCGACTTC 240  
CC |||||||  
Db 890 CCCTGCTGGAGTACCGGACGACGACGACCTTGTGCAACCAACCGCGAGATCCCGACTTC 949  
QY 241 TTCAAGAGAGCTTCCCGAGGCGCTTCACTGGAGAGAACCAACCTTACGAGAGCGGC 300  
CC |||||||  
Db 950 TTCAAGAGAGCTTCCCGAGGCGCTTCACTGGAGAGAACCAACCTTACGAGAGCGGC 1009  
QY 301 GGCATCTGTACCGGCCCAACGAGACACCAAGCTTGGAGGCAACTGTATCTTCAAGGTG 360  
CC |||||||  
Db 1010 GGCATCTGTACCGGCCCAACGAGACACCAAGCTTGGAGGCAACTGTATCTTCAAGGTG 1069  
QY 361 AAGTGACGCGCACCACTTCCCGCGGACGCGCCCGTGATGAAGAACAGAGCGCGGC 420  
CC |||||||  
Db 1070 AAGTGCTGGGACCACTTCCCGCGGACGCGCCCGTGATGAAGAACAGAGCGCGGC 1129  
QY 421 TGGAGCCCGACCGAGGTGTATACCCCGAAGAGGGCGTGTGCGGCGCGGACGTG 480  
CC |||||||  
Db 1130 TGGAGCCCGACCGAGGTGTATACCCCGAAGAGGGCGTGTGCGGCGCGGACGTG 1189  
QY 481 ATGGCCTGAAGGTGGGCGACCGGCACTGATCTGCCACCACTACACAGCTACCGGAGC 540  
CC |||||||  
Db 1190 ATGGCCTGAAGGTGGGCGACCGGCGGTGATCTGCCACCACTACACAGCTACCGGAGC 1249  
QY 541 AAGAGGCCGTGCGCGCCTGACCATGCGCGCTTCCACTTACCGACATCCGGCTCCAG 600  
CC |||||||  
Db 1250 AAGAGGCCGTGCGCGCCTGACCATGCGCGCTTCCACTTACCGACATCCGGCTCCAG 1309  
QY 601 ATGCTCGGGAAGAAGAGGACGAGTACTTGGAGCTGTACGAGGCCAGCGTGCGCCGGTAC 660  
CC |||||||  
Db 1310 ATGCTCGGGAAGAAGAGGACGAGTACTTGGAGCTGTACGAGGCCAGCGTGCGCCGGTAC 1369  
QY 661 AGCGACCTGCGCGGAGAGGCCCACTGA 687  
CC |||||||  
Db 1370 AGCGACCTGCGCGGAGAGGCCCACTGA 1396

RESULT 5  
ABL41175  
ID ABL41175 standard; DNA; 1424 BP.

XX ABL41175;  
AC

XX 12-AUG-2002 (first entry)  
DT

DE Cr-449-tandem-actin fusion protein nucleotide sequence.

XX Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;  
KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;  
KW fluorescence resonance energy transfer; fusion protein; ds.  
XX  
OS Synthetic.



OS Heteractis crispa.  
XX  
FH Key Location/Qualifiers  
FT CDS 14..1423  
FT /\*tag= a  
FT /product= "Cr-449-tandem-actin fusion protein"  
FT /note= "the stop codon is not indicated"  
FT misc\_feature 698..709  
FT /\*tag= b  
FT /note= "linker nucleotide fragment"  
FT misc\_feature 1394..1405  
FT /\*tag= c  
FT misc\_feature 1406..1423  
FT /\*tag= d  
FT /note= "partial actin fragment"  
XX  
PN WO200230965-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 12-OCT-2001; 2001WO-US032080.  
XX  
PR 12-OCT-2000; 2000US-0240018P.  
PR 16-JUL-2001; 2001US-0306131P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;  
XX  
DR WPI; 2002-444170/47.  
DR P-PSDB; ABB07997.  
XX  
XX  
PT Novel nucleic acid encoding Stichodactylidaen chromoprotein and its  
PT fluorescent mutant useful as coloring agent, labels in analyte detection  
PT assays, markers in recombinant DNA applications and filters in  
PT sunscreens.  
XX  
XX  
PS Disclosure; Fig 13; 81pp; English.  
XX  
XX The invention relates to a nucleic acid present in other than its natural  
CC environment and encoding an Stichodactylidaen chromoprotein or its  
CC fluorescent mutant, where the fluorescent protein has an emission maximum  
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are  
CC useful in applications employing a chromo or fluorescent nucleic acid or  
CC protein. Recombinant vectors comprising the nucleic acid is useful for  
CC producing an Anthozoan chromo and/or fluorescent protein. The  
CC chromoproteins, and their fluorescent mutants are useful as colouring  
CC agents capable of imparting colour or pigment to a particular composition  
CC of matter. The chromoproteins can be incorporated into a variety of  
CC different compositions including food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g. animals and plants, and as labels in  
CC analyte detection assays, e.g. assays for biological analytes of interest  
CC (see ABL41167 for a detailed description of the various uses of the  
CC chromoproteins). The present sequence represents the H. crispa Cr-449-  
CC tandem-actin fusion protein nucleotide sequence  
XX  
SQ Sequence 1424 BP; 301 A; 477 C; 452 G; 194 T; 0 U; 0 Other;  
  
Query Match 97.4%; Score 669.4; DB 6; Length 1424;  
Best Local Similarity 98.4%; Pred. No. 1.2e-108;  
Matches 676; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
  
QY 1 ATGGTGAAGCGGCTGCTGAAGGAGATATGCGCATCAAGATGTACATGAGGGACACCGT 60  
|||  
XX 14 ATGGTGAAGCGGCTGCTGAAGGAGATGCGCATCAAGATGTACATGAGGGACACCGT 73  
|||  
QY 61 AACGGCCACTACTTCAAGTGCAGGGGCGAGGGCGACGCAACCCCTTCGCCGCGCACCG 120  
|||  
DB 74 AACGGCCACTACTTCAAGTGCAGGGGCGAGGGCGACGCAACCCCTTCGCCGCGCACCG 133  
|||  
QY 121 AGCATGAGAAATCCACGTGACCGAGGGGCCCCCTGCGCTTGCGCTTGACATCTTGCGC 180  
|||  
DB 134 AGCATGCGGATCCACGTGACCGAGGGGCCCCCTGCGCTTGCGCTTGACATCTTGCGC 193  
|||

QY 181 CCCTGCTCGAGTACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCGACTTC 240  
|||  
DB 194 CCCTGCTCGAGTACGGCAGCAGGACCTTCGTGCACCAACCGCCGAGATCCCGACTTC 253  
|||  
QY 241 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAAACCAACCTACGAGGCGGC 300  
|||  
DB 254 TTCAAGCAGAGCTTCCCGAGGGCTTCACCTGGGAGAGAAACCAACCTACGAGGCGGC 313  
|||  
QY 301 GGCATCCTGACCCGCCACCAAGACACCAAGCCTTGAGGGCACTGCCTGATCTACAAGGTG 360  
|||  
DB 314 GGCATCCTGACCCGCCACCAAGACACCAAGCCTTGAGGGCACTGCCTGATCTACAAGGTG 373  
|||  
QY 361 AAGGTGCAAGCAGCAACTTCCCGCGGACCGGCCCCCGTGATGAAGAACAGAGCGGCGGC 420  
|||  
DB 374 AAGGTGCTGGCAGCAACTTCCCGCGGACCGGCCCCCGTGATGAAGAACAGAGCGGCGGC 433  
|||  
QY 421 TGGAGCCCGACACCGAGGTGTATACCCCGAGAACGGCGTGTGCGCGCGGAACGTG 480  
|||  
DB 434 TGGAGCCCGACACCGAGGTGTATACCCCGAGAACGGCGTGTGCGCGCGGAACGTG 493  
|||  
QY 481 ATGGCCCTGAAGGTGGGCGACCGGCACCTGATCTGCCACCACTACACGACTACCGAGC 540  
|||  
DB 494 ATGGCCCTGAAGGTGGGCGACCGGCGGCTGATCTGCCACCACTACACGACTACCGAGC 553  
|||  
QY 541 AAGAAGGCCGTGCGCGCCCTTGACCATGCCCCGGCTTCCAATTACCGACATCCGGCTCCAG 600  
|||  
DB 554 AAGAAGGCCGTGCGCGCCCTTGACCATGCCCCGGCTTCCAATTACCGACATCCGGCTCCAG 613  
|||  
QY 601 ATGCTGCGGAAGAAGAAGAGAGTACTTCGAGCTGTACGAGGCCAGCGTGCGCGGTAC 660  
|||  
DB 614 ATGCTGCGGAAGAAGAAGAGAGTACTTCGAGCTGTACGAGGCCAGCGTGCGCGGTAC 673  
|||  
QY 661 AGGACCTGCCCCGAGAAAGGCCCAACTGA 687  
|||  
DB 674 AGGACCTGCCCCGAGAAAGGCCCAACAGA 700  
|||  
  
RESULT 6  
ACA61025  
ID ACA61025 standard; DNA; 1424 BP.  
XX  
XX ACA61025;  
AC  
XX  
XX 09-JUL-2003 (first entry)  
DT  
XX  
XX DNA encoding chromoprotein Cr-449-tandem-actin.  
DE  
XX  
XX Chromo/fluorescent domain; labeled fusion protein;  
KW site-specific gene modification; chromoprotein; colouring agent;  
KW food composition; pharmaceutical; cosmetic; Cr-449-tandem-actin; gene;  
KW ds.  
OS  
XX Anthozoa.  
XX  
FH Key Location/Qualifiers  
FT CDS 14..1423  
FT /\*tag= a  
FT /product= "Cr-449-tandem-actin"  
FT /partial  
FT /note= "No stop codon given"  
XX  
XX  
PN WO2003031590-A2.  
XX  
PD 17-APR-2003.  
XX  
XX  
PF 10-OCT-2002; 2002WO-US032560.  
XX  
XX  
PR 12-OCT-2001; 2001US-00976673.  
PR 11-FEB-2002; 2002US-0356225P.  
PR 22-MAY-2002; 2002US-0383336P.  
XX  
XX  
PA (CLON-) CLONTECH LAB INC.

XX  
PI Lukyanov SA;  
XX  
DR WPI; 2003-381709/36.  
DR P-PSDB; ABU09922.  
XX  
PT New nucleic acid encoding polypeptide products having at least two linked  
PT chromo/fluorescent domains, useful for generating transgenic plants or  
PT animals or site-specific gene modifications in cell lines.  
XX  
PS Disclosure; Fig 2; 68bp; English.  
XX  
CC The invention describes a nucleic acid encoding a polypeptide product  
CC comprising a first and a second chromo/fluorescent domain, optionally  
CC joined by a linking domain. The first and second chromo/fluorescent  
CC domains associate with each other under intracellular conditions so that  
CC the encoded polypeptide assumes a tertiary structure. The nucleic acid  
CC and the protein are useful in producing labeled fusion proteins that have  
CC a precise and predictable signal to fusion partner ratio. The nucleic  
CC acid may also be used in generating transgenic, non-human plants or  
CC animals or site-specific gene modifications in cell lines. The  
CC chromoproteins may be used as colouring agents, as a food composition, in  
CC pharmaceuticals or cosmetics, as labels in analyte detection assays or as  
CC selectable markers in recombinant DNA applications. This sequence encodes  
CC chromo/fluorescent domain fusion protein Cr-449-tandem-actin  
XX  
SQ Sequence 1424 BP; 301 A; 477 C; 452 G; 194 T; 0 U; 0 Other;  
  
Query Match 97.4%; Score 669.4; DB 7; Length 1424;  
Best Local Similarity 98.4%; Pred. No. 1.2e-108;  
Matches 676; Conservative 0; Mismatches 11; Indels 0; Gaps 0;  
  
QY 1 ATGGTGAGCGGCTGCTGAAGAGATATGCGCATCAAGATGTATCATGAGGCAACCGTG 60  
DB 14 ATGGTGAGCGGCTGCTGAAGAGATGCGCATCAAGATGTATCATGAGGCAACCGTG 73  
QY 61 AACGGCCACTACTTCAAGTGGAGGGGCGAGGGCGGCAACCCCTTCCGGGCAACCCAG 120  
DB 74 AACGGCCACTACTTCAAGTGGAGGGGCGAGGGCGGCAACCCCTTCCGGGCAACCCAG 133  
QY 121 AGCATGAGAATCCACGTGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 180  
DB 134 AGCATGCGGATCCACGTGACGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 193  
QY 181 CCCTGCTGCGAGTACGCGACGAGACCTTGTGACACACCGGCGAGATCCCGGACTTC 240  
DB 194 CCCTGCTGCGAGTACGCGACGAGACCTTGTGACACACCGGCGAGATCCCGGACTTC 253  
QY 241 TTCAAGCAGAGCTTCCCGGAGGGGCTTCACTGTGGAGAGAACCAACCACTTACGAGACGGC 300  
DB 254 TTCAAGCAGAGCTTCCCGGAGGGGCTTCACTGTGGAGAGAACCAACCACTTACGAGACGGC 313  
QY 301 GGCATCCTGACCGGCCCAACGAGACCAACGCTTGAAGGCAACTGCTTACTTCAAGGTG 360  
DB 314 GGCATCCTGACCGGCCCAACGAGACCAACGCTTGAAGGCAACTGCTTACTTCAAGGTG 373  
QY 361 AAGGTGACGGGCAACCTTCCCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 420  
DB 374 AAGGTGCTGGGGCAACCTTCCCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 433  
QY 421 TGGAGAGCCGACGAGGTGTGTATCCCGAGAACGCGGTGTGTGTGGGGGGGGGAAAGTGTG 480  
DB 434 TGGAGAGCCGACGAGGTGTGTATCCCGAGAACGCGGTGTGTGTGGGGGGGGGAAAGTGTG 493  
QY 481 ATGGCCCTGAAGGTGGGGGAGCCGGGACCTGATCTGCCACCACTACCAAGCTACCGGAGC 540  
DB 494 ATGGCCCTGAAGGTGGGGGAGCCGGGAGCTGATCTGCCACCACTACCAAGCTACCGGAGC 553  
QY 541 AAGAAGCCGCTGCGGCTTGACCATGCGCGGCTTCCACTTCAACCGACATCCGGCTCCAG 600  
DB 554 AAGAAGCCGCTGCGGCTTGACCATGCGCGGCTTCCACTTCAACCGACATCCGGCTCCAG 613  
QY 601 ATGCTGCGGAAGAAGAGAGAGTAATTGAGCTGTATGAGGCCAGCGTGGCCCGGTAC 660

DB 614 ATGCTGCGGAAGAGAGAGAGTAATTGAGCTGTATGAGGCCAGCGTGGCCCGGTAC 673  
QY 661 AGCGAAGCTGCGCGAGAGAGGCGCACTGA 687  
DB 674 AGCGAAGCTGCGCGAGAGAGGCGCACTGA 700  
  
RESULT 7  
ACA61026  
ID ACA61026 standard; DNA; 1376 BP.  
XX  
AC ACA61026;  
XX  
DT 09-JUL-2003 (first entry)  
XX  
DE DNA encoding chromoprotein HcRed-Cr1-tandem.  
XX  
KW Chromo/fluorescent domain; labeled fusion protein;  
KW site-specific gene modification; chromoprotein; colouring agent;  
KW food composition; pharmaceutical; cosmetic; HcRed-Cr1-tandem; gene; ds.  
XX  
OS Anthozoa.  
XX  
FH Key location/qualifiers  
FT 1..1376  
FT CDS /\*tag= a  
FT /product= "HcRed-Cr1-tandem"  
FT /transl\_except= (pos:679..680, aa:Asn)  
XX  
PN WO2003031590-A2.  
XX  
PD 17-APR-2003.  
XX  
PE 10-OCT-2002; 2002WO-US032560.  
XX  
PR 12-OCT-2001; 2001US-00976673.  
PR 11-FEB-2002; 2002US-0356225P.  
PR 22-MAY-2002; 2002US-0383336P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA;  
XX  
DR WPI; 2003-381709/36.  
DR P-PSDB; ABU09923.  
XX  
PT New nucleic acid encoding polypeptide products having at least two linked  
PT chromo/fluorescent domains, useful for generating transgenic plants or  
PT animals or site-specific gene modifications in cell lines.  
XX  
PS Disclosure; Fig 3; 68bp; English.  
XX  
CC The invention describes a nucleic acid encoding a polypeptide product  
CC comprising a first and a second chromo/fluorescent domain, optionally  
CC joined by a linking domain. The first and second chromo/fluorescent  
CC domains associate with each other under intracellular conditions so that  
CC the encoded polypeptide assumes a tertiary structure. The nucleic acid  
CC and the protein are useful in producing labeled fusion proteins that have  
CC a precise and predictable signal to fusion partner ratio. The nucleic  
CC acid may also be used in generating transgenic, non-human plants or  
CC animals or site-specific gene modifications in cell lines. The  
CC chromoproteins may be used as colouring agents, as a food composition, in  
CC pharmaceuticals or cosmetics, as labels in analyte detection assays or as  
CC selectable markers in recombinant DNA applications. This sequence encodes  
CC chromo/fluorescent domain fusion protein HcRed-Cr1-tandem  
XX  
SQ Sequence 1376 BP; 381 A; 292 C; 357 G; 346 T; 0 U; 0 Other;  
  
Query Match 65.0%; Score 446.6; DB 7; Length 1376;  
Best Local Similarity 78.2%; Pred. No. 1.4e-69;  
Matches 536; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

QY	3	GGTGAGCGGCGCTGCTGAAGAGAGATATGCGCATCAAGATGTATCATGGAGGGACCGCTGAA	62
Db	692	GATGTCGTGTTGTTGAAAGAAAGTATGCGCATCAAGATGTATCATGGAGGGACCGCTTAA	751
QY	63	CGGCCACTACTTCAAGTGGAGGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAG	122
Db	752	TGGCCATTATTTCAGTGTGAAGAGAGAGGAGACGGCAACCCATTTCAGAGTACGCAGAG	811
QY	123	CATGAGAAATCCACGTGACCGAGGGCGCCCCCTGCGCTTCGCTTCGACATCCTGGCCCC	182
Db	812	CATGAGGATTCATGTATCACCGAAGGGGCTCCATTACCATTTGCCTTCGACATTTTGGCA	871
QY	183	CTGCTGCGAGTACGGGACGACGACCTTCTGTGCAACCAACCGCCGAGATCCCGCATTT	242
Db	872	GTGTTGTGAGTACGGGACGACGACCTTTGTGCCAATACGGCAGAGATTCCCGATTTCTT	931
QY	243	CAAGCAGAGCTTCCCCCGAGGGCTTCACTGTGGAGAGAAACCACTTACGAGGACGGCGG	302
Db	932	CAAGCAGTCTTCCCTGAAGGCTTACTTGGAAAGAACCAACCACTATGAGATGAGAGG	991
QY	303	CATCCTGACCGCCCAACGACACACAGCCTGGAGGGCACTGCTGATCTACAAGGTGAA	362
Db	992	CATTCTTACTGCTCATCAGACACAAGCCTGAGAGGGAACTGCCTTATATACAAGGTGAA	1051
QY	363	GGTGACAGGACCAACTTCCCCCGCGACGCGCCCCGCTGATGAAGAACAAGAGCGGCGCTG	422
Db	1052	AGTCCATGGTATACCAATTTTCTGCTGATGCGCCCGTGATGAAGAACAATCAGAGAGATG	1111
QY	423	GGAAGCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGTGTGCGGCCGGAACGTGAT	482
Db	1112	GGAAGCCAGCACGTGAGGTGGTTATCCAGAGAAATGGTGTCTGTGTGACGTAATGTGAT	1171
QY	483	GGCCCTGAAGGTGGGGCGACCGGCACTGTATCTGCCACCACTACACAGCTACCGGACAA	542
Db	1172	GGCCCTTAAAGTCGGTATCGTCAATTGATCTGCCATCACTAATCTTCTTACAGTCCAA	1231
QY	543	GAAAGCCGTGCGCGCCCTGACCATGCCCCGGCTTCCACTTCAACCGACATCCGGCTCCAGT	602
Db	1232	GAAAGCAGTCCGTGCTTGAACAATGCCAGGATTTTCAATTTAACAGACATCCGCTTCAGAT	1291
QY	603	GCTGCGAAGAAGAAGACGAGTACTTCGAGCTGTACAGGCCACGCTGCGCCGTTACAG	662
Db	1292	GCTGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCACTCTGTGGCTAGGTACAG	1351
QY	663	CGACCTGCCCCGAGAGGCCAACTGA	687
Db	1352	TGATCTTCTGAAAAAGCAAAATTGA	1376

RESULT 8	
ABL41170	
ID	ABL41170 standard; cDNA; 681 BP.
XX	
AC	ABL41170;
XX	
DT	12-AUG-2002 (first entry)
XX	
DE	H. crispae fluorescent protein mutant 44-9 encoding cDNA.
XX	
KW	Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRET;
KW	colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;
KW	fluorescence resonance energy transfer; gene expression; mutant; ss.
XX	
OS	Heteractis crispae.
XX	
FH	
FT	Key location/Qualifiers
CT	1..681
FT	/*tag= a
FT	/product= "mutant fluorescent protein"
XX	
PN	WO200230965-A2.
XX	
PD	18-APR-2002.

XX 12-OCT-2001; 2001WO-US032080.  
PF 12-OCT-2000; 2000US-0240018P.  
XX 16-JUL-2001; 2001US-0306131P.  
PR  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;  
XX  
DR WPI; 2002-444170/47.  
DR P-PSDB; ABB07992.  
XX  
XX Novel nucleic acid encoding Stichodactylidaen chromoprotein and its  
PT fluorescent mutant useful as coloring agent, labels in analyte detection  
PT assays, markers in recombinant DNA applications and filters in  
PT sunscreens.  
XX  
PS Claim 5; Fig 6; 81pp; English.

The invention relates to a nucleic acid present in other than its natural environment and encoding an Stichodactylidae chromoprotein or its fluorescent mutant, where the fluorescent protein has an emission maximum ranging from 580-660 nm. The polynucleotides and encoded proteins are useful in applications employing a chromo or fluorescent nucleic acid or protein. Recombinant vectors comprising the nucleic acid is useful for producing an Anthozoan chromo and/or fluorescent protein. The chromoproteins, and their fluorescent mutants are useful as colouring agents capable of imparting colour or pigment to a particular composition of matter. The chromoproteins can be incorporated into a variety of different compositions including food compositions, pharmaceuticals, cosmetics, living organisms, e.g. animals and plants, and as labels in analyte detection assays, e.g. assays for biological analyses of interest (see AB41167 for a detailed description of the various uses of the chromoproteins). The present sequence represents the H. crispa fluorescent protein mutant 44-9 encoding cDNA

Query Match	64.2%;	Score 441.2;	DB 6;	Length 681;
Best Local Similarity	78.2%;	Pred. No. 1.3e-68;		
Matches 530; Conservative	0;	Mismatches 148;	Indels 0;	Gaps 0

[illegible]

Db 424 AGCACTGAGTGGTTTATCCAGAAATGTCCTGTGTGACGTAATGTATGCGCCCTT 483  
Qy 490 AAGGTGGCGCAGCCGGCACCCTGATCTGCCACCACTACACCAAGTACCGAGCAAGAAGGCC 549  
Db 484 AAAGTCGGTATCGTGTGTTGATCTGCCATCACTATCTTTACAGGTCCAGAAAGCA 543  
Qy 550 GTGCGCGCCCTGACCATGCCCGGCTTCCACTTCAACGACATCCGGCTCCAGATGCTGCGG 609  
Db 544 GTCCGTCCTTGACAATGCCAGGATTTTCATTTTACAGACATCCGCTTCAAGATGCTGAGG 603  
Qy 610 AAGAAGAAAGACGAGTACTTGCAGCTGTACGAGCCGACGCGGCCGTTACAGCCGACCTG 669  
Db 604 AAAGAGAAAGACGAGTACTTGTGAAGTGTACGAAAGCATCTGTGGCTAGGTACAGTATCTT 663  
Qy 670 CCCGAGAGGCCCACTGA 687  
Db 664 CCTGAAAAAGCAAAATTGA 681

## RESULT 9

ABL41171

ID ABL41171 standard; cDNA; 681 BP.

XX ABL41171;

DT 12-AUG-2002 (first entry)

DE H. crispa fluorescent protein mutant 44-6 encoding cDNA.

KW Stichodactylidaen; chromoprotein; fluorescent; anthozoan; food; FRFT;  
KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;  
KW fluorescence resonance energy transfer; gene expression; mutant; ss.

XX Heteractis crispa.

XX Key Location/Qualifiers  
FT CDS 1.681  
FT /\*tag= a  
FT /product= "mutant fluorescent protein"

XX WO200230965-A2.

XX 18-APR-2002.

XX 12-OCT-2001; 2001WO-US032080.

XX 12-OCT-2000; 2000US-0240018P.

XX 16-JUL-2001; 2001US-0306131P.

XX (CLON-) CLONTECH LAB INC.

XX Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;

XX WPI; 2002-444170/47.

XX P-PSDB; ABB07993.

PT Novel nucleic acid encoding Stichodactylidaen chromoprotein and its  
PT fluorescent mutant useful as coloring agent, labels in analyte detection  
PT assays, markers in recombinant DNA applications and filters in  
PT sunscreens.

XX Claim 5; Fig 8; 81pp; English.

XX The invention relates to a nucleic acid present in other than its natural  
CC environment and encoding an Stichodactylidaen chromoprotein or its  
CC fluorescent mutant, where the fluorescent protein has an emission maximum  
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are  
CC useful in applications employing a chromo or fluorescent nucleic acid or  
CC protein. Recombinant vectors comprising the nucleic acid is useful for  
CC producing an Anthozoan chromo and/or fluorescent protein. The  
CC chromoproteins, and their fluorescent mutants are useful as colouring  
CC agents capable of imparting colour or pigment to a particular composition  
CC of matter. The chromoproteins can be incorporated into a variety of

CC different compositions including food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g. animals and plants, and as labels in  
CC analyte detection assays, e.g. assays for biological analytes of interest  
CC (see ABL41180 for a detailed description of the various uses of the  
CC chromoproteins). The present sequence represents the H. crispa  
CC fluorescent protein mutant 44-6 encoding cDNA

XX Sequence 681 BP; 185 A; 145 C; 178 G; 173 T; 0 U; 0 Other;

SQ Query Match 64.0%; Score 439.6; DB 6; Length 681;

Best Local Similarity 78.0%; Pred. No. 2.4e-68;

Matches 529; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

Qy 10 GGCCCTGCTGAAGAGAGATATGCGCATCAAGATGTACATGAGAGGCAACCGTGAACGGCCAC 69  
Db 4 GGTGTGTTGAAAGAAAGTATGCGCATCAAGATGTACATGAGAGGCAACCGTGAATGGCCAT 63

Qy 70 TACTTCAAGTGCAGAGGGCGAGGGCGAGCAACCCCTTCCGCGGACCCAGACATGAGA 129  
Db 64 TATTTCAGTGTGAAGAGAGAGGAGACGGAACCCATTGTCAGGTACGACAGCATGAGG 123

Qy 130 ATCCACGTGACCGAGGGGCGCCCTTGCCTTGCCTTGCATCTCTGCCCCCTGCTGC 189  
Db 124 ATTCATGTCAACGAAAGGGGCTCCATTACCATTTGCTTGCACATTTTGGCACCGGTGTGT 183

Qy 190 GAGTACGGGACGAGGACCTTCTGTGCACCAACCGCGAGATCCCGACTTCTTCAAGCAG 249  
Db 184 GCGTACGGGACGAGGACCTTGTGTCCACCAATACGGCAGAGATTCGCCGATTTCTTCAAGCAG 243

Qy 250 AGCTTCCCGAGGGGCTTCACTTGGGAGAAACCAACCACTTACGAGAGCGGGGCTCTCTG 309  
Db 244 TCTTCCCTGAAGGCTTACTTGGGAAAGAACCAACCACTTACGAGATGAGGCTCTCTT 303

Qy 310 ACCGCCACAGAGACACAGCCTTGAGGGCAACTGCTGATCTTCAAGGTGAAGGTGCAC 369  
Db 304 ACTGCTCATCAGAGACAAAGCCTTGAGGGGAACTGCTTATATACAAAGGTGAAGTCTCTT 363

Qy 370 GGCACCACTTCCCGCGAGCGCCCGGTGATGAAGAAAGAGCGGCGCTGGAGCCC 429  
Db 364 GGTACCAATTTTCTGCTGATGCGCCCGGTGATGAAGAAAGATCAGAGATGGAGCCA 423

Qy 430 AGCACCGAGGTGTGTATCCCGAGAAACGCGTGTGCGGCGCGGAACGTATGCGCCCTG 489  
Db 424 AGCACTGAGGTGTGTATTCAGAGAAATGTGTCTGTGTGACGTAATGTATGCGCCCTT 483

Qy 490 AAGGTGGGCGACCGGACCTGATCTGCGCACCACTACACAGCTACCGGACGAAGAGGCC 549  
Db 484 AAAGTCGGTATCGTGTGTGTATCTGCGCATCACTATCTTCTTACAGGTCCAAAGAGCA 543

Qy 550 GTGCGCGCCCTGACCAATGCGCGGCTTCACTTCAACGACATCCGGCTCCAGATGCTGCGG 609  
Db 544 GTCCGTGCTTGAACAATGCCAGATTTTATTTTACAGACATCCGCTTCAAGATGCTGAGG 603

Qy 610 AAGAAGAAAGACGAGTACTTGCAGCTGTACAGAGCGGCGGTGAGGAGGAGGAGGAGGAGG 669  
Db 604 AAAGAGAAAGACGAGTACTTGAACGTGTACGAAGCATCTGTGCTAGGTACAGTATCTT 663

Qy 670 CCCGAGAGGCCCAACTGA 687  
Db 664 CCTGAAAAAGCAAAATTGA 681

## RESULT 10

ABL41180

ID ABL41180 standard; cDNA; 684 BP.

XX ABL41180;

DT 12-AUG-2002 (first entry)

DE H. crispa chromoprotein mutant hccp mut C148S encoding cDNA.

KW Stichodactylidaen; chromoprotein; fluorescent; anthozoan; food; FRFT;



KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;  
KW fluorescence resonance energy transfer; gene expression; mutant; ss.  
XX  
OS Heteractis crispa.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..684  
FT /\*tag= a  
FT /product= "mutant fluorescent protein"  
XX  
PN WO200230965-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 12-OCT-2001; 2001WO-US032080.  
XX  
PR 12-OCT-2000; 2000US-0240018P.  
PR 16-JUL-2001; 2001US-0306131P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;  
XX  
DR WPI; 2002-444170/47.  
DR P-PSDB; ABB07998.  
XX  
XX  
PT Novel nucleic acid encoding Stichodactyliiden chromoprotein and its  
PT fluorescent mutant useful as coloring agent, labels in analyte detection  
PT assays, markers in recombinant DNA applications and filters in  
PT sunscreens.  
XX  
PS Example; Fig 15; 81pp; English.  
XX  
CC The invention relates to a nucleic acid present in other than its natural  
CC environment and encoding an Stichodactyliiden chromoprotein or its  
CC fluorescent mutant, where the fluorescent protein has an emission maximum  
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are  
CC useful in applications employing a chromo or fluorescent nucleic acid or  
CC protein. Recombinant vectors comprising the nucleic acid is useful for  
CC producing an Anthozoan chromo and/or fluorescent protein. The  
CC chromoproteins, and their fluorescent mutants are useful as colouring  
CC agents capable of imparting colour or pigment to a particular composition  
CC of matter. The chromoproteins can be incorporated into a variety of  
CC different compositions including food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g. animals and plants, and as labels in  
CC analyte detection assays, e.g. assays for biological analytes of interest  
CC (see ABL41167 for a detailed description of the various uses of the  
CC chromoproteins). The present sequence represents the cDNA encoding an  
CC alternative embodiment of the H. crispa fluorescent protein mutant C148S  
XX  
SQ Sequence 684 BP; 188 A; 145 C; 178 G; 173 T; 0 U; 0 Other;  
  
Query Match 63.8%; Score 438.2; DB 6; Length 684;  
Best Local Similarity 77.6%; Pred. No. 4.3e-68;  
Matches 530; Conservative 0; Mismatches 153; Indels 0; Gaps 0;  
  
QY 5 TGAGCGGCTGCTGAAGAGATGCGCATCAAGATGTACATGAGGGCACCGTGAACG 64  
Db 2 TGGCTGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACATGGAAGGCACGGTTATG 61  
QY 65 GCCACTACTTCAAGTGCAGGGGCGAGGGCGACGCAACCCCTTCGCCGGCACCCAGACA 124  
Db 62 GCCATTATTTCAGTGTGAAGAGAGAGGAGACGCGCAACCACTTTACAGTACGACAGCA 121  
QY 125 TGAAGATCCACGTGACCGAGGGGCCCCCTGCCCCCTTGCCTTCGACATCCTGGCCCCCT 184  
Db 122 TGAGGATTTCATGTCAACCGAAGGGGCTCCATTACCATTTGCCCTTCGACATTTGGCACCGT 181  
QY 185 GCTGCGAGTACGGCAGCAGACCTTTCGTGCACCAACCCCGAGAGATCCCGACTTCTTCA 244  
Db 182 GTTGTGAGTACGGCAGCAGACCTTTCGTGCACCAATACGGCAGAGATTCCCGATTCTTCA 241  
QY 245 AGCAGAGCTTCCCCGAGGGCTTCACCTGGGAGAGAACCAACCACTTACGAGAGCGCGGCA 304

Db 242 AGCAGTCTTCCCTGAAGCCTTTACTTGGAAAGAACCAACCTATGAAAGATGAGCGCA 301  
QY 305 TCCTGACCGCCACCAAGACACAGCCTGAGGGCACTGCTGATCTACAAGGTGAAGG 364  
Db 302 TTCTTACTGCTCATCAGACACACAGCCTGAGGGGAACTGCCTTATATACAAGGTGAAG 361  
QY 365 TGCAACGGCACCACTTCCCGCCCGACGGCCCCCGTGAATGAAGAACAGAGCGCGCTGGG 424  
Db 362 TCCTTGTTACCAATTTTCTCTGCTGATGGCCCCCGTGAATGAAGAACATCAGAGGATGGG 421  
QY 425 AGCCCAACACCGAGGTGTGTACCCCGAGAACGGCGTGTGCGGCCGGAACGTGATGG 484  
Db 422 AGCCAAACACTGAGGTGTGTTATCCAGAGAAATGTGTCTGTGTGACGTTAATGTGATGG 481  
QY 485 CCCTGAAAGGTGGCGCACCGGCACTGATCTGCGCACTACACCACTACCGGAGCAAGA 544  
Db 482 CCTTAAAGTCGGTGATCGTGTGTTGATCTGCGCATCTCTAATCTTACAGGTCCAAGA 541  
QY 545 AGGCGGTGCGCGCCCTGACCATGCCCGGCTTCCACTTCAACGACATCCGGCTCCAAGTGC 604  
Db 542 AAGCAGTCCGTGCTTGAACATGCCAGGATTTCAATTACAGACATCCGCTTCAGATGC 601  
QY 605 TGCGGAAGAAGAGACGAGTACTTCGAGCTGTACAGGCCACGCGTGCCCGTACAGCG 664  
Db 602 CGAGGAAAAAGAAAGACGAGTACTTTGAACGTGTACGAAGCATCTGTGGCTAAGTACAGTG 661  
QY 665 ACCTGCCCCGAGGAAGGCCAATGCA 687  
Db 662 ATCTTCTGAAAAAGCAAAATTGA 684  
  
RESULT 11  
ABL41169  
ID ABL41169 standard; cDNA; 684 BP.  
XX  
AC ABL41169;  
XX  
DT 12-AUG-2002 (first entry)  
XX  
DE H. crispa fluorescent protein mutant C148S encoding cDNA.  
XX  
KW Stichodactyliiden; chromoprotein; fluorescent; anthozoan; food; FRET;  
KW colouring agent; pharmaceutical; cosmetic; immunoassay; biosensor; gene;  
KW fluorescence resonance energy transfer; gene expression; mutant; ss.  
XX  
OS Heteractis crispa.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..684  
FT /\*tag= a  
FT /product= "mutant fluorescent protein"  
XX  
PN WO200230965-A2.  
XX  
PD 18-APR-2002.  
XX  
PF 12-OCT-2001; 2001WO-US032080.  
XX  
PR 12-OCT-2000; 2000US-0240018P.  
PR 16-JUL-2001; 2001US-0306131P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;  
XX  
DR WPI; 2002-444170/47.  
DR P-PSDB; ABB07991.  
XX  
XX  
PT Novel nucleic acid encoding Stichodactyliiden chromoprotein and its  
PT fluorescent mutant useful as coloring agent, labels in analyte detection  
PT assays, markers in recombinant DNA applications and filters in  
PT sunscreens.

XX Claim 5; Fig 4; 81pp; English.  
PS  
CC The invention relates to a nucleic acid present in other than its natural  
CC environment and encoding an Stichodactylidaen chromoprotein or its  
CC fluorescent mutant, where the fluorescent protein has an emission maximum  
CC ranging from 580-660 nm. The polynucleotides and encoded proteins are  
CC useful in applications employing a chromo or fluorescent nucleic acid or  
CC protein. Recombinant vectors comprising the nucleic acid is useful for  
CC producing an Anthozoan chromo and/or fluorescent protein. The  
CC chromoproteins, and their fluorescent mutants are useful as colouring  
CC agents capable of imparting colour or pigment to a particular composition  
CC of matter. The chromoproteins can be incorporated into a variety of  
CC different compositions including food compositions, pharmaceuticals,  
CC cosmetics, living organisms, e.g. animals and plants, and as labels in  
CC analyte detection assays, e.g. assays for biological analyses of interest  
CC (see ABL41167 for a detailed description of the various uses of the  
CC chromoproteins). The present sequence represents the H. crlspa  
CC fluorescent protein mutant C1485 encoding cDNA  
XX  
SQ Sequence 684 BP; 187 A; 146 C; 178 G; 173 T; 0 U; 0 Other;

Query Match 63.6%; Score 436.6; DB 6; Length 684;  
Best Local Similarity 77.5%; Pred. No. 8.2e-68;  
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 5 TGAGCGGCTGCTGAAGAGATATGCCATCAAGATGTACATGAGGGCACCGTGAACG 64  
Db 2 TGGCTGGTTTGTGAAGAAGATATGCCATCAAGATGTACATGGAAGGCACGGTTAATG 61  
QY 65 GCCACTACTTCAAGTGCAGAGGGGAGAGGGGCGGCAACCCCTTCGCCGGGACCCAGAGCA 124  
Db 62 GCCATTATTTCAGTGTGAAGAGAGAGGAGACGGCAACCCATTTCACAGGTACGAGAGCA 121  
QY 125 TGAGATCCAGTGAACCGAGGGCGCCCCCTTGCCCTTCGACATCTGGCCCCCT 184  
Db 122 TGAGATTCATGTCAACCGAGGGGCTCCATTACCATTTGCCCTTCGACATTTGGCACCGT 181  
QY 185 GCTGCGAGTACGGGAGAGACCTTCTGTGACACCAACCGCGAGATCCCCGACTTCTTCA 244  
Db 182 GTTGTAGTACGGGAGAGACCTTGTCCACCATACGGGAGAGATTCGCCATTTCTTCA 241  
QY 245 AGCAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCACTACGAGAGCGGCGCA 304  
Db 242 AGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACCACTATGAGATGAGGCA 301  
QY 305 TCCTGACCGCCACCAGAGACCAAGCTGAGGGGCAACTGCTGATCTACAGGTGAAGG 364  
Db 302 TTCTTACTGCTCATCAGAGACAAAGCTGAGGGGAACTGCTTATATACAAAGGTGAAG 361  
QY 365 TGCACGGCACCACTTCCCGCGGAGCGCCCCGTGATGAAGAACAAGCGCGGCTGGG 424  
Db 362 TCCTTGGTACCAATTTCTGCTGATGGCCCGTGATGAAGAACAATCAGAGGATGGG 421  
QY 425 AGCCACGACCGAGGTGTGTAACCCGAGAACGGCGTGTGCGGGCGCGAAGCTGATGG 484  
Db 422 AGCCAAAGCACTGAGGTGTTATCCAGAGAAATGTTCTGTGTGACGTAATGTGATGG 481  
QY 485 CCCTGAAGGTGGGCGACCGGCACTGATCTGCCACCACTACACAGCTACCGGAGCAAGA 544  
Db 482 CCCTTAAGTCGGTGTGTCGTTGATGCTCATCTCTATATCTTCTTACAGGTCCAAGA 541  
QY 545 AGCCGCTGGCGCCCTGACCATGCCCCGCTTCCACTTACCCGACATCCGGCTCCAGATGC 604  
Db 542 AAGCAGTCCGTGCTTGACAATGCCAGATTTCAATTACAGACATCCGCTTTCAGATGC 601  
QY 605 TGCAGAAAGAAAGAGACGATCTTGAAGTGTACGAGGCGACGCTGGCCGGTACAGCG 664  
Db 602 CGAGGAAACGAAAGACGAGTACTTTGAAGTGTACGAAGCATCTGTGCTAGGTACAGTGT 661  
QY 665 ACCTGCCGAGAAAGGCCAAGTGA 687  
Db 662 ATCTTCCTGAAGAAAGCAAAATTGA 684

RESULT 12  
ABA00802  
ID ABA00802 standard; cDNA; 760 BP.  
XX  
AC ABA00802;  
XX  
DT 01-APR-2003 (first entry)  
XX  
DE Wild type chromoprotein cDNA.  
XX  
KW Gene; kindling fluorescent protein; kindling stimulus; movement;  
KW labeling; fluorescence resonance energy transfer; FRET;  
KW bioluminescence resonance energy transfer; BRET; biosensor;  
KW automated screening; ss.  
XX  
OS Heteractis crispa.  
XX  
FH Key location/Qualifiers  
FT CDS 77..760  
FT /\*tag= a  
FT /product= "Chromoprotein"  
XX  
FN WO200296924-A1.  
XX  
PD 05-DEC-2002.  
XX  
PF 24-MAY-2002; 2002WO-US016379.  
XX  
PR 25-MAY-2001; 2001US-0293752P.  
PR 11-OCT-2001; 2001US-0329176P.  
XX  
PA (CLON-) CLONTECH LAB INC.  
XX  
PI Lukyanov SA, Chudakov D, Lukyanov K;  
XX  
DR WPI; 2003-156788/15.  
DR P-PSDB; AAG79765.  
XX  
PT Novel nucleic acid that is present in other than its natural environment  
PT and that encodes kindling fluorescent protein, is useful in labeling  
PT protocols, e.g. labeling proteins, organelles, cells and organisms.  
XX  
PS Example; Fig 4; 96pp; English.  
XX

CC The sequences given in ABA00799-805 encode wild type and mutant kindling  
CC fluorescent proteins. The proteins go from a first substantially non-  
CC fluorescent or non-fluorescent state to a second fluorescent state upon  
CC exposure to a kindling stimulus. The kindling proteins are useful for  
CC detecting an entity such as a protein, organelle or cell in a composition  
CC such as a cell or a multicellular composition (preferably a multicellular  
CC organism), by providing the entity as an entity labeled with the kindling  
CC protein, kindling the kindling fluorescent protein label with a kindling  
CC stimulus to produce a kindled kindling fluorescent protein label, and  
CC exciting the kindled kindling fluorescent protein label with light and  
CC detecting any fluorescence from it to detect the entity. The method  
CC monitors the movement of the entity. The fluorescent proteins and the  
CC cDNA encoding them are useful in labeling protocols, e.g., labeling  
CC proteins, organelles, cells and organisms, as biological labels or  
CC markers, in protein labeling and tagging applications. The fluorescent  
CC kindling proteins are useful as detectable labels, as labels in analyte  
CC detection assays, in fluorescence resonance energy transfer (FRET)  
CC applications, in bioluminescence resonance energy transfer (BRET)  
CC applications, as biosensors in prokaryotic and eukaryotic cells, in  
CC applications involving the automated screening of arrays of cells  
CC expressing fluorescent reporting groups, in high through-put screening  
CC assays, as second messenger detectors, and in fluorescent activated cell  
CC sorting assays  
XX

SQ Sequence 760 BP; 209 A; 166 C; 189 G; 196 T; 0 U; 0 Other;  
Query Match 63.6%; Score 436.6; DB 7; Length 760;

Best Local Similarity 77.5%; Pred. No. 8.2e-68;			
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;			
QY	5	TGAGCGCCTGCTGAAGAGAGATGCGCATCAAGATGTACATGAGGGCA	CGGTGAACG 64
Db	78	TGGCTGGTTTGTGAAGAAGATATGGCATCAAGATGTACATGAGAGCA	CGGTTAATG 137
QY	65	GCCACTACTTCAAGTGCAGGGGCGAGGGCGGACCGCAACCCCTTCG	CGCCGCAACCAAGCA 124
Db	138	GCCATTATTCAAGTGTGAAGAGAGAGGAGACGGCAACCCATTTA	CAGGTAACGACAGCA 197
QY	125	TGAGAAATCCACGTGACCGAGGGGCGCCCCCTGCCCCCTTCGCTT	CGACATCCTGGCCCCCT 184
Db	198	TGAGGATTCAATGTCAACCGAAGGGGCTCCATTACATTGGCCTT	CGACATTTGGCACCGT 257
QY	185	GCTGCGAGTACGGCAGCAGACCTTTCGTGACACACACCGCGAGAT	CCCCGACTTCTTCA 244
Db	258	GTTGTGAGTACGGCAGCAGACCTTTCGTGACACCATACGACAGAT	TCCGATTTCTTCA 317
QY	245	AGCAGAGCTTCCCCGAGGCTTACCTGGAGAGAAACCACTACGAG	AGAGACGGCGCA 304
Db	318	AGCAGTCTTCCCTGAAGGCTTACTTGGGAAGAACCACCACTATGA	AGATGAGAGCA 377
QY	305	TCCTGACCGCCCAACCAAGACACAGCCTGGAGGGCACTGCCTGAT	CTACAAGGTGAAG 364
Db	378	TTCTTACTGCTCATCAGGACACACAGCCTGAGGGGAATGCTCTTA	TATACAAGGTGAAG 437
QY	365	TGCACGGCAACCACTTCCCCCGCGAGCGCCCGTGATGAAGACA	AGAGCGGCGTGGG 424
Db	438	TCCTTGATCAATTTTCTGCTGATGCGCCCGTGATGAAGACAAT	CAGAGGATGGG 497
QY	425	AGCCCAACCGAGGTGTGTACCCGAGAAACGGCGCTGTGCGGCCG	GGAACGTGATGG 484
Db	498	AGCCATGCACTGAGGTGTTTATCCAGAGAATGTCCTGTGTGGA	CGTAATGTGATGG 557
QY	485	CCCTGAAGTGGGCGACCGGACCTGATCTGCCACACTACCACTAC	CACTACCGAGCAAG 544
Db	558	CCCTTAAGTCGGTGATCGTCTTGATCTGCCATCTCTATACTTCT	TACAGAGTCCAAGA 617
QY	545	AGGCCGTGCGCGCCCTGACCATGCGCGCTTCCACTTACCCGACAT	CCGGCTCCAGATGC 604
Db	618	AAGCAGTCCGTGCTTGACAATGCCAGGATTTCAATTTACAGACAT	CCGCCCTTCAGATGC 677
QY	605	TGCGAAGAAGAGAGAGTACTTGTGAGCTGTACGAGGCCAGCGT	GGCCGGTACAGCG 664
Db	678	CGAGGAAAAAGAAAGACGAGTACTTGAACCTGTACGAAGCATCT	GCTAGTACAGTG 737
QY	665	ACCTGCGCGAGAGGCCCACTGA	687
Db	738	ATCTTCTGAAAAAGCAATTGA	760
RESULT 13			
ABA00806			
ID	ABA00806 standard; cDNA; 760 BP.		
XX	ABA00806;		
AC			
XX	01-APR-2003 (first entry)		
DT			
XX	Multiple mutant Chromoprotein #2 cDNA.		
DE			
XX			
KW	Gene; kindling fluorescent protein; kindling stimulus; movement;		
KW	labeling; fluorescence resonance energy transfer; FRET;		
KW	bioluminescence resonance energy transfer; BRET; biosensor;		
KW	automated screening; ss.		
OS	Heteractis cribrpa.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	77..760	
FT	/*tag= a		
FT	/product= "Multiple mutant Chromoprotein #2"		

FT	mutation	replace(182,A)
FT		/*tag= c
FT	mutation	replace(503,T)
FT		/*tag= m
FT	mutation	replace(549..550, AT)
FT		/*tag= n
FT	mutation	replace(594,T)
FT		/*tag= h
FT	mutation	replace(662..663, AT)
FT		/*tag= j
FT	mutation	replace(678,C)
FT		/*tag= o
FT	mutation	replace(686,A)
FT		/*tag= 1
XX		
PN	WO200296924-A1.	
XX		
PD	05-DEC-2002.	
XX		
PF	24-MAY-2002; 2002WO-US016379.	
XX		
PR	25-MAY-2001; 2001US-0293752P.	
PR	11-OCT-2001; 2001US-0329176P.	
XX		
PA	(CLON-) CLONTECH LAB INC.	
XX		
PI	Lukyanov SA, Chudakov D, Lukyanov K;	
XX		
DR	WPI; 2003-156788/15.	
DR	P-PSDB; AAG79769.	
XX		
PT	Novel nucleic acid that is present in other than its natural environment	
PT	and that encodes kindling fluorescent protein, is useful in labeling	
PT	protocols, e.g. labeling proteins, organelles, cells and organisms.	
XX		
PS	Example; Fig 8; 96pp; English.	
XX		
CC	The sequences given in ABA00799-805 encode wild type and mutant kindling	
CC	fluorescent proteins. The proteins go from a first substantially non-	
CC	fluorescent or non-fluorescent state to a second fluorescent state upon	
CC	exposure to a kindling stimulus. The kindling proteins are useful for	
CC	detecting an entity such as a protein, organelle or cell in a composition	
CC	such as a cell or a multicellular composition (preferably a multicellular	
CC	organism), by providing the entity as an entity labeled with the kindling	
CC	protein, kindling the kindling fluorescent protein label with a kindling	
CC	stimulus to produce a kindled kindling fluorescent protein label, and	
CC	exciting the kindled kindling fluorescent protein label with light and	
CC	detecting any fluorescence from it to detect the entity. The method	
CC	monitors the movement of the entity. The fluorescent proteins and the	
CC	cDNA encoding them are useful in labeling protocols, e.g., labeling	
CC	proteins, organelles, cells and organisms, as biological labels or	
CC	markers, in protein labeling or tagging applications. The fluorescent	
CC	kindling proteins are useful as detectable labels, as labels in analyte	
CC	detection assays, in fluorescence resonance energy transfer (FRET)	
CC	applications, in bioluminescence resonance energy transfer (BRET)	
CC	applications, as biosensors in prokaryotic and eukaryotic cells, in	
CC	expressing fluorescent reporting groups, in high through-put screening	
CC	assays, as second messenger detectors, and in fluorescent activated cell	
CC	sorting assays	
XX		
SQ	Sequence 760 BP; 208 A; 167 C; 193 G; 192 T; 0 U; 0 Other;	
Query Match		
Best Local Similarity 63.6%; Score 436.6; DB 7; Length 760;		
Matches 529; Conservative 0; Mismatches 154; Indels 0; Gaps 0;		
QY	5	TGAGCGCCTGCTGAAGAGAGATGCGCATCAAGATGTACATGAGGGCACCGTGAACG 64
Db	78	TGGCTGGTTTGTGAAGAAGATATGGCATCAAGATGTACATGGAAGGACGGTTAATG 137
QY	65	GCCACTACTTCAAGTGCAGGGGCGAGGGCGGACCGCAACCCCTTCGCGGCAACCAAGCA 124



Db	138	GCCATTATTTCAAGTGTGAAGAGAGGGAGACGGCAACCATTTGCAGGTACGCAGAGCA	197		
QY	125	TGAGATTCACGTGACCGAGGGGCGCCCTGCTTCGCTTGACATCTGGCCCT	184		
Db	198	TGAGATTTCATGTACCGAAGGGGCTTCATTACCATTTGCTTGACATTTTGGACCGT	257		
QY	185	GCTGCGAGTACGGCAGCAGGACCTTCGTGCACACACCGCCGAGATCCCGACTTCTTCA	244		
Db	258	GTTGTAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGAGATTCCCGATTCTTCA	317		
QY	245	AGCAGAGCTTCCCGAGGGCTTACCTGGAGAGAAACCAACCTACGAGGACGGCGCA	304		
Db	318	AGCAGTCTTTCCTGAAGGCTTTACTTGGGAAAGAACCAACCTATGAAGATGAGGCA	377		
QY	305	TCCTGACCGCCACCAGGACACCAGCCTGGAGGGCACTGCCTGATCTACAAGTGAAG	364		
Db	378	TTCCTTACTGCTCATCAGGACACAAGCCTGAGGGGAACTGCCTTATATACAAGTGAAAG	437		
QY	365	TGCACGGCACCACTTCCCGCGCAGCGGCCCGTGATGAAGAACAAGCGCGGCTGGG	424		
Db	438	TCCTTGATCAATTTTCTGCTGATGGCCCGTGATGAAGAACAAATCAGAGGATGGG	497		
QY	425	AGCCGACACCGAGGTGTGTACCCCGAGAACGGCGTGTGCGCGCGGAACGTGATGG	484		
Db	498	AGCCAAAGCACTGAGGTGTTTATCCAGAGAAATGTTCTGTGTGACGTAGCGTGATGG	557		
QY	485	CCCTGAAGGTGGGCGGACCGGCACTGTATGCCACCACTACACAGCTAACCGGACAAGA	544		
Db	558	CCCTTAAAGTCGGTGTATCTGTTGATCTGCCATCACTATATCTTCTTACAGGTCCAAGA	617		
QY	545	AGGCCGTGCGGCGCCCTGACCATGCCCCGCTTCCACTTACCCGACATCCGGCTCCAGATGC	604		
Db	618	AAGCAGTCCGTGCTTGAACATGCCAGGATTTCAATTTCACAGACCAACCGCTTCAAGATGC	677		
QY	605	TGCGGAGAGAAGAGACGAGTACTTCGAGCTGTACGAGGCCAGCGTGCGCCGGTACAGCG	664		
Db	678	GGAGGAAAGAGAAAGACGAGTACTTTGAACCTGTACGAAGCATCTGTGGCTAAGTACAGTG	737		
QY	665	ACCTGCCCGAGAGAGGCCCACTGA	687		
Db	738	ATCTTCCCTGAAAAAGCAAAATTGA	760		
RESULT 14					
ABL41167					
ID	ABL41167 standard; cDNA; 910 BP.				
XX					
AC	ABL41167;				
XX					
DT	12-AUG-2002 (first entry)				
XX					
DE	H. crisp a chromoprotein wild-type base isoform hCFP640 cDNA sequence.				
XX					
KW	Stichodactylidaen; chromoprotein; fluorescent; anthozoan; food; FRET;				
KW	colouring agent; pharmaceutical; cosmetic; immunassay; biosensor; gene;				
KW	fluorescence resonance energy transfer; gene expression; hCFP640; ss.				
XX					
OS	Heteractis crisp a.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	82..765			
FT		/*tag= a			
FT		/product= "chromoprotein"			
XX					
PN	WO200230965-A2.				
XX					
PD	18-APR-2002.				
XX					
PF	12-OCT-2001; 2001WO-US032080.				
XX					
PR	12-OCT-2000; 2000US-0240018P.				
PR	16-JUL-2001; 2001US-0306131P.				
XX					

PA	(CLON-) CLONTECH LAB INC.			
XX				
PI	Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;			
XX				
DR	WPI; 2002-444170/47.			
DR	P-PSDB; ABB07989.			
XX				
PT	Novel nucleic acid encoding Stichodactylidaen chromoprotein and its			
PT	fluorescent mutant useful as coloring agent, labels in analyte detection			
PT	assays, markers in recombinant DNA applications and filters in			
PT	sunscreens.			
PS	Claim 5; Fig 1; 81pp; English.			
XX				
CC	The invention relates to a nucleic acid present in other than its natural			
CC	environment and encoding an Stichodactylidaen chromoprotein or its			
CC	fluorescent mutant, where the fluorescent protein has an emission maximum			
CC	ranging from 580-660 nm. The polynucleotides and encoded proteins are			
CC	useful in applications employing a chromo or fluorescent nucleic acid or			
CC	protein. Recombinant vectors comprising the nucleic acid is useful for			
CC	producing an Anthozoan chromo and/or fluorescent protein. The			
CC	chromoproteins, and their fluorescent mutants are useful as colouring			
CC	agents capable of imparting colour or pigment to a particular composition			
CC	of matter. The chromoproteins can be incorporated into a variety of			
CC	different compositions including food compositions, pharmaceuticals,			
CC	cosmetics, living organisms, e.g. animals and plants, and as labels in			
CC	analyte detection assays, e.g. assays for biological analytes of			
CC	interest. The chromoproteins may be incorporated into adducts with			
CC	analyte specific antibodies or their binding fragments and subsequently			
CC	employed in immunoassays for analytes of interest in a complex sample.			
CC	They are also useful as selectable markers in recombinant DNA			
CC	applications, e.g. the production of transgenic cells and organisms, in			
CC	suncreens, as selective filters, and in fluorescence resonance energy			
CC	transfer (FRET) applications, where the proteins serve as donor and/or			
CC	acceptors in combination with a second fluorescent protein or dye, e.g. a			
CC	fluorescent protein. The proteins also find use as biosensors in			
CC	prokaryotic and eukaryotic cells, e.g. as Ca2+ ion indicator, as pH			
CC	indicator, as phosphorylation indicator, as an indicator of other ions,			
CC	e.g. magnesium, sodium, potassium, chloride and halides and in			
CC	applications involving the automated screening of arrays of cells			
CC	expressing fluorescent reporting groups by using microscopic imaging and			
CC	electronic analysis. Screening can be used for drug discovery and in the			
CC	field of functional genomics e.g. where the subject proteins are used as			
CC	markers of whole cells to detect changes in multicellular reorganisation			
CC	and migration, e.g. formation of multicellular tubules by endothelial			
CC	cells, migration of cells, wound healing and neurite outgrowth. The			
CC	fluorescent protein also finds use in high throughput screening assays,			
CC	in fluorescence activated cell sorting applications, as a label to mark a			
CC	population of cells, as in vivo marker in animals, in protease cleavage			
CC	assays, in assays to determine the phospholipid composition in biological			
CC	membranes and as a fluorescent timer, in which the switch of one			
CC	aging of the fluorescent protein is used to determine the activation/			
CC	deactivation of gene expression, e.g. developmental gene expression, cell			
CC	cycle dependent gene expression or circadian rhythm specific gene			
CC	expression. The present sequence represents the H. crisp a wild-type base			
CC	isoform hCFP640 (hccp) cDNA encoding a chromoprotein			
XX				
SQ	Sequence 910 BP; 250 A; 199 C; 212 G; 249 T; 0 U; 0 Other;			
QY	Query Match	63.3%;	Score 435;	DB 6; Length 910;
Db	Best Local Similarity	77.3%;	Pred. No. 1.6e-67;	
QY	Matches 528; Conservative	0;	Mismatches 155;	Indels 0; Gaps 0;
QY	5	TGAGCGGCTGCTGAAGAGAGATATGCGCATCAAGATGTACATGGAGGCGACCGTGAACG	64	
Db	83	TGGCTGTTGTGTGAAGAAAGTATGCGCATCAAGATGTACATGGAGGCGACCGTTAATG	142	
QY	65	GCCACTACTTCAAGTGCAGAGGCGAGGCGCAACCCCTTGCCCGGACCCAGAGCA	124	
Db	143	GCCATTATTTCAAGTGTGAAGAGAGAGGAGCGCAACCCATTTTACAGGTACGAGAGCA	202	
QY	125	TGAGAAATCCAGTGACCGAGGGCGCCCTGCTTGCCTTGCACATCTGGCCCCCT	184	



Db 203 TGAGGATTCATGTCAACCGAGGGGCTCCATTACCATTTGCCCTTCGACATTTTGGCACCGT 262

QY 185 GCTGCGAGTACGGCAGCAGACCTTGTGCACCACACACCGCGAGATCCCGACTTCTTCA 244

Db 263 GTTGTGAGTACGGCAGCAGACCTTGTGCCACCATACGGCAGAGATCCCGATTTCTTCA 322

QY 245 AGCAGAGCTTCCCGAGGGCTTCACTGGGAGAGAACCACTACGAGACGGCGGCA 304

Db 323 AGCAGTCTTCCCTGAAGGCTTACTTGGGAAAGAACCAACCTATGAAGATGGAGCA 382

QY 305 TCCTGACCGCCACCAGACACCAAGCTTGAGGGCAACTGCCTGATCTACAAGTGAAG 364

Db 383 TTCTTACTGCTCATCAGACACAAAGCTTGAGGGGAAGTGCCTTATATACAAGTGAAG 442

QY 365 TGCAGCGCACCAACTTCCCGCGAGCGCCCGTGATGAAGAACAAAGCGCGGCTGGG 424

Db 443 TCCTTGTAACCAATTTTCTGCTGTGATGGCCCCGTGATGAAGAACAAATCAGAGATGGG 502

QY 425 AGCCAGCACCGAGGTGTGTACCCCGAGAACGCGGTGCTGTGGCGCGGAACGTGATGG 484

Db 503 AGCCATGCACTGAGGTGTTATCCAGAGAATGTGTCTGTGTGACGTAATGTGATGG 562

QY 485 CCCTGAAGGTGGGGCAGCCGGCAGCTGATCTGCCACCACTACACCAAGCTACCGAGCAGA 544

Db 563 CCCTTAAAGTGGGTGATCGTCGTTGATCTGCCATCTTATACTTCTTACAGGTCCAGA 622

QY 545 AGCCGCTGCGCGCCCTGACCATCGCGGCTTCCACTTCACCGACATCCGGCTCCAGATGC 604

Db 623 AAGCAGTCCGTGCTTGACAATGCAAGATTTCATTTTACAGACATCCGCTTCAGATGC 682

QY 605 TGCAGAAAGAAAGAGCAGTACTTGAGCTGTACGAGGCCAGCGTGGCCCGGTACAGCG 664

Db 683 CGAGGAAACGAAAGACGAGTACTTGAACCTGTACGAAGCACTGTGGCTAGGTACAGTG 742

QY 665 ACCTGCGCGAGAGGCCAACTGA 687

Db 743 ATCTTCTGAAAAAGCAATTTGA 765

RESULT 15

ABL41182

ID ABL41182 standard; cDNA; 910 BP.

XX AC ABL41182;

XX 12-AUG-2002 (first entry)

DT 12-AUG-2002 (first entry)

XX H. crispa chromoprotein wild-type base isoform hccp cDNA.

XX Stichodactylidae; chromoprotein; fluorescent; anthozoan; food; FRRT;

KW colouring agent; pharmaceutical; cosmetic; immunosens; biosensor; gene;

KW fluorescence resonance energy transfer; gene expression; hcfP640; ss.

XX Heteractis crispa.

OS

XX Key location/Qualifiers

FT CDS 82..765

FT /\*tag= a

FT /product= "chromoprotein"

XX

PN WO200230965-A2.

XX 18-APR-2002.

PD

XX 12-OCT-2001; 2001WO-US032080.

PF

XX 12-OCT-2000; 2000US-0240018P.

PR 16-JUL-2001; 2001US-0306131P.

XX

XX (CLON-) CLONTECH LAB INC.

PA

XX Lukyanov SA, Fradkov AF, Lukyanov KA, Gurskaya NG;

XX WPI; 2002-444170/47.

DR P-PSDB; ABB08000.

XX Novel nucleic acid encoding Stichodactylidae chromoprotein and its

PT fluorescent mutant useful as coloring agent. Labels in analyte detection

PT assays, markers in recombinant DNA applications and filters in

PT sunscreens.

XX Example; Fig 17; 81pp; English.

PS

XX The invention relates to a nucleic acid present in other than its natural

CC environment and encoding an Stichodactylidae chromoprotein or its

CC fluorescent mutant, where the fluorescent protein has an emission maximum

CC ranging from 580-660 nm. The polynucleotides and encoded proteins are

CC useful in applications employing a chromo or fluorescent nucleic acid or

CC protein. Recombinant vectors comprising the nucleic acid is useful for

CC producing an Anthozoan chromo and/or fluorescent protein. The

CC chromoproteins, and their fluorescent mutants are useful as colouring

CC agents capable of imparting colour or pigment to a particular composition

CC of matter. The chromoproteins can be incorporated into a variety of

CC different compositions including food compositions, pharmaceuticals,

CC cosmetics, living organisms, e.g. animals and plants, and as labels in

CC analyte detection assays, e.g. assays for biological analytes of interest

CC (see ABL41167 for a detailed description of the various uses of the

CC chromoproteins). The present sequence represents the cDNA encoding an

CC alternative embodiment of the H. crispa wild-type base isoform hccp cDNA

XX

SQ Sequence 910 BP; 250 A; 199 C; 212 G; 249 T; 0 U; 0 Other;

Query Match 63.3%; Score 435; DB 6; Length 910;

Best Local Similarity 77.3%; Pred. No. 1.6e-67;

Matches 528; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

QY 5 TGAAGCGGCTGCTGAAGAGAGTATGCCGATCAAGATGTACATGTGAGGGGACCGTGAACG 64

Db 83 TGCGTGGTTGTGAAAGAAAGTATGCCGATCAAGATGTACATGGAAGGACCGTTAATG 142

QY 65 GCCACTACTTCAAGTGGAGGGGGCGAGGGCGGCAACCCCTTGCCTTGACATCTGCCCCCT 124

Db 143 GCCATTATTTCAGTGTGAAGAGAGAGGAGAGCGGCAACCCATTACAGGTACGAGAGCA 202

QY 125 TGAGATCCACGTGACGAGGGGGCCCCCTGCTTGCCTTGACATCTGCCCCCT 184

Db 203 TGAGGATTCATGTCAACGAAGGGGCTCCATTACCAATTGCTTGACATTTTGGCACCGT 262

QY 185 GCTGCGAGTACGGCAGAGGACCTTGTGACACACCGCGAGATCCCCGACTTCTTCA 244

Db 263 GTTGTAGTACGGCAGCAGACCTTGTGCCACCATACGGCAGAGATTTCCGATTTCTTCA 322

QY 245 AGCAGAGCTTCCCGAGGGCTTACCTGGGAGAGAACCAACCACTACGAGACGGCGGCA 304

Db 323 AGCAGTCTTCCCTGAAGGCTTACTTGGGAAAGAACCAACCTATGAAGTGAAGGCA 382

QY 305 TCCTGACCGCCACACAGACACCAAGCTTGAGGGCACTGCCTGATCTACAAGTGAAG 364

Db 383 TTCTTACTGCTCATCAGACACAAAGCTTGAGGGGAAGTGCCTTATATACAAGTGAAG 442

QY 365 TGCAGCGCACCAACTTCCCGCGAGCGCCCGTGATGAAGAACAAAGCGCGGCTGGG 424

Db 443 TCCTTGTAACCAATTTTCTGCTGTGATGGCCCCGTGATGAAGAACAAATCAGAGATGGG 502

QY 425 AGCCAGCACCGAGGTGTGTACCCCGAGAACGCGGTGCTGTGGCGCGGAACGTGATGG 484

Db 503 AGCCATGCACTGAGGTGTTATCCAGAGAATGTGTCTGTGACGTAATGTGATGG 562

QY 485 CCCTGAAGGTGGGGCAGCCGGCAGCTGATCTGCCACCACTACACCAAGCTACCGGAGCAAG 544

Db 563 CCCTTAAAGTGGGTGATCGTGTGATCTGCCATCTCTATACTTCTTACAGGTCCAAGA 622

QY 545 AGCCGCTGCGCGCCCTGACCATGCCCCGCTTCCACTTCAACGACATCCGGCTCCAGATGC 604

Db 623 AAGCAGTCCGTGCTTGACAATGCCAGGATTTCAATTTTACAGACATCCGCTTCAGATGC 682

Qy	605	TC	CG	GA	GA	GA	GA	CG	AG	TA	CT	TC	GA	CT	GT	AC	GA	GC	CA	GC	CG	CG	CG	TA	CA	GC	664	
Db	683	CG	AG	GA	AA	AC	GA	AG	CG	AG	TA	CT	TT	GA	AC	TG	TAC	GA	GC	AT	CT	GT	GG	CT	AG	TA	CAGTG	742
Qy	665	AC	CT	GC	CC	GA	GA	AG	GC	CA	CT	GA	687															
Db	743	AT	CT	TC	CT	GA	AA	AA	AG	CA	AA	TT	GA	765														

Search completed: July 15, 2004, 12:12:28  
Job time : 399 secs